AD_____

Award Number: DAMD17-94-J-4380

TITLE: Diagnostic Strategies for Breast Cancer: Optimizing the Tradeoffs

PRINCIPAL INVESTIGATOR: John B. Wong, M.D.

CONTRACTING ORGANIZATION: American College of Physicians
Philadelphia, Pennsylvania 19106-1572

REPORT DATE: May 1999

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1999	Final (1 Sep 94 - 30 Apr 99)

4. TITLE AND SUBTITLE
Diagnostic Strategies for Breast Cancer: Optimizing the Tradeoffs

5. FUNDING NUMBERS
DAMD17-94-J-4380

6. AUTHOR(S)
John B. Wong, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
American College of Physicians-
American Society of Internal Medicine
Philadelphia, Pennsylvania 19106-1572

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This is the final report for the "Diagnostic Strategies for Breast Cancer" grant. From over 6000 citations, we have abstracted and analyzed 85 articles involving over 27,000 patients, fine needle aspirates and breast biopsies for the evaluation of breast lesions. Our meta-analysis, the largest performed to date, suggests that fine needle aspiration has a sensitivity and specificity that ranges from 79.3% to 92.4% and from 85.7% to 99.1% depending on whether results for suspicious or atypical cytology are classified as "positive." Our decision analysis comparing fine needle aspiration to open biopsy differs from nearly all prior analyses by considering the long-term costs of care for breast cancer and the economic and clinical effects of a possible delay in diagnosis resulting from false negative cytology results. Our results suggest that open biopsy may be cost-effective because of cost savings and lives saved from fewer false negative cytology results. The likelihood of breast cancer, the likelihood of a delayed diagnosis resulting in a more advanced stage of breast cancer all affected the results.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Meta-analysis, decision analysis, breast cancer, fine needle aspiration cytology, cost-effectiveness analysis	84
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

JBW Where copyrighted material is quoted, permission has been obtained to use such material.

JBW Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

JBW Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

NA JBW For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

John B Wu 5/27/95

PI - Signature Date

Table Of Contents

Introduction

The purposes of this grant were to:

I. Perform a comprehensive structured literature review of the diagnostic tests for the evaluation of suspected breast cancer;
II. Conduct focus groups of physicians and patients to give this work expert review and feedback;
III. Construct a decision analysis evaluating the optimal diagnostic test strategy for breast cancer evaluation when comparing fine needle aspiration to open biopsy;
IV. Compare the marginal cost-effectiveness of open biopsy versus fine needle aspiration cytology taking into consideration the long-term costs and morbidity of false positives and false negatives.

Despite advances in treatment and earlier detection, breast cancer remains the leading site for newly developed cancers in women. It is the second leading cause of cancer deaths in women in the United States, affecting one in eight women from birth to death. Evaluation has evolved from a one-stage procedure that involved a breast biopsy with a frozen section while the patient was under anesthesia to determine whether the patient would awaken with a mastectomy, to a two-stage procedure which provides an opportunity for discussion of alternative breast conservation therapies. As a result, some one million breast biopsies may occur annually.[1] Fine needle aspiration cytology provides a less invasive alternative method, but the tissue sample is smaller, resulting in false positive and false negative diagnoses. False positive diagnoses lead to unnecessary breast biopsies which cause anxiety and incur economic costs. Also, false negative cytological diagnoses may result in a delayed diagnosis with potential worsening in the stage of the breast cancer. The first goal of this study was to estimate the sensitivity and specificity of fine needle aspiration by performing a comprehensive literature review. The second goal of this study was to conduct focus groups of physicians and patients with experience in the evaluation of breast lesions to guide the development of the meta-analysis and decision analysis. Although a detailed examination of tru-cut breast biopsy was beyond the scope of this grant, using results from our meta-analysis, our study goes on to develop a decision analysis as part of the third goal to estimate the life expectancy consequences of choosing fine needle aspiration versus open biopsy. Lastly, our fourth goal was to compare the cost-effectiveness of open biopsy versus fine needle aspiration taking into consideration the costs and the effectiveness of false negative and false positive cytology results from fine needle aspiration.

Body

I. Meta-analysis

Methods

During the first year of the project, then Principal Investigator, Anthony So, MD, performed part of the meta-analysis at the American College of Physicians. Dr. So and his colleagues performed an extensive preliminary meta-analysis with creation of a data entry form, selection of articles, data abstraction and preliminary analysis. They were, however, unable to complete their analysis. When we took over the grant in 1997, our study consulted physicians with expertise in the evaluation of breast lesions to give the work expert feedback. Although our initial intent was to simply update Dr. So's database, through these discussions, we found it necessary to modify the data abstraction form. Consequently, we repeated all of the steps performed in year one by Dr. So in addition to updating the literature review. We added other search terms, revised the data abstraction form and performed the meta-analysis anew. For details of Dr. So's extensive work done in the firs year of the project, refer to the year one progress report (see enclosure). The remainder of this report will refer to the work performed at New England Medical Center which was an extension of Dr. So's initial work, but, in essence, performed all of the year one activities anew along with the year two activities over the past 18 months.

Selection of Articles

A MEDLINE search was performed of all English language studies of human beings published from 1966 through 1998. We applied a previously published search list that had the highest sensitivity for detecting relevant articles along with Dr. So's prior search which used the MESH terms: BIOPSY, NEEDLE and BREAST or BREAST DISEASES, title terms FINE NEEDLE, ASPIRATION, CYTOLOGY and BREAST or MAMMARY. The article selection process involved review of the title list and exclusion of those articles with clearly inappropriate content. Only English language documents involving human subjects were included. Letters were excluded. Titles were then scanned for relevancy. The winnowed list was then subjected to a review of their abstracts. The remaining articles were then obtained and abstracted. The following exclusion criteria were then applied: 1) duplicate publication; 2) N<50; 3) absence of primary data, 4) special population; 5) absent reference or gold standard; and 6) inadequate detail to allow determination of sensitivity or specificity. When the first exclusion criteria was met, the article was excluded. Bibliographies of selected articles were examined for additional studies not discovered by the MEDLINE search.

Data Abstraction

Although a data abstraction form was created during year one of the study, upon review, additional elements were added to the data abstraction form to capture study, population and technique characteristics, which examine variation in practice and explore potential relationships between those variations and cytology results (See Appendix I). Based on our prior experience with meta-analysis of non-invasive evaluation of heart disease, we also added data fields to capture the department from which the article arose. To avoid

duplicate inclusion of data, we also added a field to capture the institution and years that the study was performed. To standardize our data collection and to permit analysis of alternative cutpoints for defining "positive" cytology results, we included categories of malignant, suspicious, atypical, benign and inadequate. We separately analyzed sensitivity and specificity hypothetically assuming that cytology was considered "positive": 1) only for patients with malignant cytological results; 2) those with malignant or suspicious cytological results; 3) those with malignant, suspicious or atypical cytological results. Some patients who undergo fine needle aspiration are classified as "inadequate" because of insufficient cellular material. Clinically, these patients would likely undergo breast biopsy to clarify the diagnosis. We therefore repeated the above meta-analyses by treating patients with an inadequate aspiration as being "positive."

Data Entry and Software

Data were entered into a Lotus 123 spreadsheet program. Once the evidence tables were complete, data were exported to a text file for analysis with SAS for Windows, version 6.1. The meta-analysis was performed using the FREQ module with the Mantel-Haentzel Chi-square.

Technical Details

We combined study results using the Mantel-Haentzel technique which excludes between study variation and may underestimate slightly the uncertainty surrounding the results. To test for homogeneity, we applied the Pearson Chi-square test (with degrees of freedom equal to one less than the number of studies). Because of the insensitivity of this test, p-values less than or equal to 0.1 were considered positive. For the purposes of this study, we did not examine verification bias and only included studies with histologic confirmation of the fine needle aspirate results.

Results

Table 1 [2-83] summarizes the overall results of the meta-analysis (Appendix II). Most studies were from sites outside of the United States conducted by pathologists and surgeons. Studies involved an average of 369 patients seen between 1981 and 1985 with a mean age of 51. Technique when reported involved most often a 10 or 20 ml syringe with a 20 to 23 gauge needle and a fixative with a PAP stain.

Table 1. Summary of Study Characteristics

Parameter	Mean (n)
Source of Publication	% (number of studies)
Department	
Family Practice	1 (1)
Gynecology	5 (4)
Oncology	5 (4)
Pathology	72 (61)
Radiation oncology	2 (2)
Radiology	24 (20)
Surgery	61 (52)
Not specified	7 (6)
United States	38 (32)
Technical details in study	% (number of studies)
Syringe (ml)	
3	2 (2)
5	6 (5)
10	25 (21)
20	34 (29)
30	2 (2)
50	1 (1)
Not specified	35 (30)
Needle (gauge)	
18	4 (3)
20	13 (11)
21	25 (21)
22	42 (36)
23	18 (15)
25	1 (1)
Not specified	25 (20)
Centrifuged	16 (14)
Slide preparation	
Air	29 (26)
Fixative	65 (55)
Not specified	20 (17)
Stain	
Diff-Quik	8 (7)
Giemsa[1]	24 (20)
Hematoxylin-Eosin	6 (5)
Other[2]	2 (2)
PAP	58 (49)
Not specified	22 (19)
Patient and Study Characteristics	Mean (number of studies)
Year study	
Began	1981 (68)
Ended	1985 (67)
# of patients	481 (59)
# of men	3 (29)
Unspecified	369 (84)
# of biopsies	367 (85)
Mean age (yrs)	
Overall	51 (21)
Youngest	23 (24)
Oldest	84 (24)

[1] May-Grunwald, Leishman

[2] Romanovsky or Liu

Table 2 resents the summary of the results of the pooling. As the cutoff criteria moves from defining a positive cytology as only those results with malignant cytology to defining a positive cytology as having either malignant, suspicious or atypical cytology, sensitivity increases but specificity declines, as expected, for mammography directed, ultrasound directed or undirected fine needle aspiration. Table 3 presents similar results but includes patients with inadequate cytology and considers those results as "positive." As mentioned above, this classification is supported by clinical practice because such results require further evaluation. By definition, inclusion of these inadequate results improves sensitivity but similarly decreases specificity.

Table 2. Summary of Test Characteristics

	Sensitivity	Specificity	P-value
Undirected	N=11,665	N=13,561	
Positive if			
Malignant	79.3	99.1	<0.001
Malignant or Suspicious	90.1	93.3	<0.001
Malignant or Suspicious or Atypical	92.4	85.7	<0.001
Directed by Mammography	N=659	N=828	
Positive if			
Malignant	65.7	99.4	<0.001
Malignant or Suspicious	80.7	92.8	<0.001
Malignant or Suspicious or Atypical	87.1	81.3	<0.001
Directed by Ultrasonography	N=761	N=433	
Positive if			
Malignant	76.5	98.6	<0.001
Malignant or Suspicious	87.8	88.5	<0.001
Malignant or Suspicious or Atypical	95.8	73.9	<0.001

Table 3. Summary of Test Characteristics including Inadequate Cytology as Positive

	Sensitivity	Specificity	p-value
Undirected	N=12,241	N=15,427	
Positive if			
Malignant	80.3	87.1	<0.001
Malignant or Suspicious	90.6	82.0	<0.001
Malignant or Suspicious or Atypical	92.8	75.9	<0.001
Directed by Mammography	N=715	N=1045	
Positive if			
Malignant	68.4	78.8	<0.001
Malignant or Suspicious	82.2	73.5	<0.001
Malignant or Suspicious or Atypical	88.1	64.4	<0.001
Directed by Ultrasonography	N=997	N=783	
Positive if			
Malignant	82.1	54.5	<0.001
Malignant or Suspicious	90.7	48.9	<0.001
Malignant or Suspicious or Atypical	96.8	40.9	<0.001

II. Focus Groups

During the first year of the study, Dr. So conducted over nine interviews with expert physicians involved in the field of breast cancer. Their results are presented in the year one progress report. As outlined in our proposal, when we took over the grant at the New England Medical Center, instead of performing patient focus groups, we consulted local experts with experience in breast cancer regarding their opinions in the role of fine needle aspiration and breast biopsy. These physicians included radiologists, cytopathologists, surgeons and medical oncologists as well as social workers who are primarily located in the Breast Health Center. These physicians were consulted about their practice patterns, sources of variability in reported outcomes and special considerations regarding fine needle aspiration. Out of these discussions, we added data fields to the data abstraction form that captured technical details such as gauge needle used and ml syringe used. These conversations also formed the basis for our short-term quality of life disutility estimates for undergoing fine needle aspiration or open biopsy.

III. Decision Analysis

Decision Model

We considered four alternative strategies. They were: 1) fine needle aspirate (FNA) defining positives as those with malignant cytology; 2) FNA defining positives as those with malignant or suspicious cytology; 3) FNA defining positives as those with malignant, suspicious or atypical cytology; and 4) initial open biopsy. Patients may or may not have breast cancer and the results from these procedures may be true or false positives or negatives. Those with either true or false "positive" cytology results then undergo open biopsy as would occur clinically. To estimate the subsequent prognosis for these patients, we constructed a simple 3 state Markov model. The states of health included 1) those who are well with a benign breast lesion who do not have breast cancer, 2) those who have breast cancer and 3) a dead state of health. The computer simulation follows a hypothetical cohort of 10,000 identical women who move through these states of health over time. Time is modeled as a one year cycle, during which time, some members of the cohort may die. The simulation tracks all individuals crediting those alive in any given year for their survival and for their cost of care. By following all 10,000 identical patients until all have died, the simulation estimates the average life expectancy and lifetime costs for each strategy (see below).

Breast Cancer Survival

Average survival times were based on 5-year relative survival rates according to stage using 1986-1993 SEER data (Table 4). These relative survivals were converted to annual excess mortality rates using the Declining Exponential Approximation for Life Expectancy (DEALE).[84] The subsequent overall life expectancy for the cohort was estimated with a Markov model (see above) by including death other causes (as occurs within the general population) using an additive mortality model.[85, 86]

Table 4. SEER survival data and Costs of breast cancer by stage from Kaiser Permanente data (also see Table 8)

Stage	Initial 6-month Costs Inflation Adjusted[3] ($1998)	Annual Continuing Care Costs Inflation Adjusted ($1998)	Terminal Care Costs Inflation Adjusted ($1998)	5-year Survival (all races) %	Annual Excess Mortality Rate from Breast Cancer %	Life Exp (All races) in Years
Local	16,779	2,160	21,866	96.8	0.65	25.8
Regional	21,173	2,669	21,866	75.9	5.12	11.4

Life Exp = Life expectancy in years

[3] Using the medical care cost component of the Consumer Price Index from 1992 to 1998

Delay in Diagnosis

Sensitivity and specificity estimates for fine needle aspiration were based on the meta-analysis performed in goal one of this grant. Open biopsy was assumed to have perfect sensitivity and specificity. Patients with false negative cytology may experience a delay in the ultimate diagnosis of their breast cancer because of the false reassurance provided by falsely negative fine needle aspirate cytology. The effect of this delay in diagnostic staging of the disease is not known. Two studies involving 39 patients in total suggest that a false negative cytology resulted in more than a 3 month delay in the ultimate diagnosis of the underlying malignancy for 15 of these patients (38%).[87, 88] For our analysis, we assumed that 50% of these patients (19% of all patients with false negative results) might progress to a more advanced stage of disease, i.e., from local to regional disease because of the falsely negative cytological result. Regional disease results in a substantially decreased survival (Table 4) and higher treatment costs of care (see below).

Quality of Life Estimates

Table 5 summarizes the quality of life estimates used in our analysis.[89, 90] Long-term quality of life was taken from published studies regarding the patient and public perception of quality of life following breast surgery. Short-term disutilities are subtracted from the overall quality-adjusted survival and are based on discussions with physicians familiar with the care of patients with breast cancer.

Table 5. Utility Values

Variable	Value
Long-term quality adjustment factor Breast cancer[4]	0.85
Short-term morbidity quality adjustment factor[5] Fine needle aspiration Open breast biopsy	 -1 day -1 month

IV. Cost-effectiveness Analysis

Economic Costs

We performed a MEDLINE search of all English language studies of economic studies in breast cancer published from 1966 through 1998, using the search terms costs, mammography and breast cancer. The data sources for economic estimates included in the studies were reviewed and critiqued. The results of the critique and the rationale for the costs used are presented below. Although we could have used local variable costs for

[4] Each year that a patient with breast cancer survives is credited for living 0.85 quality-adjusted life years to take into consideration morbidity and uncertainty related to the disease

[5] Patients undergoing each procedure have this utility deducted from their overall quality-adjusted survival to reflect the morbidity related to undergoing the procedure. All patients with "positive" cytology undergo subsequent open biopsy.

these procedures as suggested in the year one progress report, we instead applied median cost estimates from national average reimbursement data which are more generalizable.[91] Two fundamentally different approaches have been used to estimate the costs associated with breast cancer. Some studies assigned charges attributable to breast cancer by subtracting the average costs associated with the care of a sample of comparable women without breast cancer. Others identified all service use associated with the diagnosis of breast cancer and then assigned costs to each service and summed the costs for all services.[92, 93] One study used the latter approach but adjusted for care received for conditions unrelated to breast cancer by subtracting the costs of care received by average patients without the cancer.[94]

To estimate lifetime direct medical expenses attributable to breast cancer, many studies separate treatment costs into 3 categories: initial therapy (the first 3-6 months after diagnosis), continuing care, and terminal care (the last 6 months of life). Baker et al[95] (Table 6) calculated the lifetime costs based on an average survival time of 10 years for women diagnosed with breast cancer (total $) but did not estimate costs according to stage of diagnosis or patient age. Subsequent analyses by Eddy[96] included stage-specific data but did not include the effects of age or comorbid conditions. Furthermore, Eddy relied on Medicare data to estimate costs, which may not be generalizable to younger women (Table 7). This study estimated a total cost (in 1984 dollars) of $36,926 for breast cancer (initial therapy $6,859; maintenance $21,409; and terminal care $8,658).

Taplin et al estimated the total and net costs of medical care for breast cancer according to stage, age, and comorbidity.[97] Net costs from the Group Health Cooperative of Puget Sound were calculated as the difference between the costs of care of women with breast cancer and the average costs of care for female enrollees without breast cancer, matched according to age. Differences in costs by stage of diagnosis, age, and comorbidity were separately evaluated using multivariate regression analysis.

Based on our review of the literature, we used estimates from Kaiser Permanente (Tables 4 and 8).[94] Costs attributable to breast cancer were derived by subtracting from the costs of each cancer patient the cost rate among health plan members of the same age (in 5-year intervals) and sex. We used their initial, interim and terminal costs within the Markov model for all patients who were alive. Those who died from breast cancer incurred the terminal care costs.

Table 6. Average Charges in 1984 Dollars for Breast Cancer (Baker)

Medicare plan	Initial 3 months ($)	Continuing Care ($)	Terminal 6 months ($)
Hospital insurance			
Inpatient	5,730	184	9,256
Skilled nursing facility	80	18	791
Home health agency	26	6	176
Supplemental medical			
Physicians	634	95	1687
Outpatient	114	22	27
Home health agency	4	1	49
Other	1,018	157	2,912
Total	7,606	483	15,137

Table 7. Breast Cancer Cost Data by Stage (Eddy)

Cost of Initial Treatment	Cost ($)
DCIS (0)	5,559
Stage I	5,880
Stage II	6,150
Stage III	6,549
Stage IV	6,863
Continuing care per month	239
Terminal care for breast cancer	14,053
Terminal care for other causes	10,814

Table 8. Costs of Care for Breast Cancer Patients at Kaiser Permanente in 1992 dollars

Stage	Initial care for 6 months ($)	Continuing Care, per quarter ($)	Terminal Care for 6 months ($)
CIS	8,515	888	11,222
Local	10,835	958	14,962
Regional	12,273	1,423	20,323
Distant	NA	2,921	20,610
Unknown	NA	1,308	18,630
Age			
35-49	11,791	1,078	28,196
50-64	11,159	991	21,426
65-79	10,054	1,104	16,587
>=80	9,135	1,353	9,937

NA = Not available

Table 9 summarizes procedure related costs based on median reimbursable physician fees in 1998 [91] for performing the procedure and interpreting the results along with the additional cost for directing the aspiration or biopsy with mammography or ultrasonography in comparison to other previously published data.[98-102]

V. Results

Decision Analysis and Cost-effectiveness Analysis

Table 10 presents the results of our analysis. Because of the decrease in life expectancy and the increased cost of care for advanced disease, our base case suggests that open biopsy may be cost-effective when compared to fine needle aspiration. Future savings offset its higher initial cost. The results are consistent with the inclinations of the physician focus groups to pursue breast biopsy when cytology results are "malignant, suspicious or atypical."

Sensitivity Analysis

The results, however, were sensitive to variation in the underlying variable estimates. For example, if the pretest probability of breast cancer fell below 18% (baseline 46%, based on the prevalence of breast cancer in the meta-analysis), then open biopsy would cost more than $50,000 per DQALY gained compared to core biopsy. If the sensitivity of fine needle aspiration exceeded 94% (baseline 92.4%) then open biopsy would again have a cost-effectiveness ratio exceeding $50,000 per DQALY gained. If the delay in diagnosis from a false negative fine needle aspirate resulted in 14% or fewer patients (baseline 19%) subsequently presenting with an advanced stage of disease, then again, fine needle aspiration would be preferable. Doubling the cost of open biopsy raised its cost-effectiveness ratio to $41,400 per DQALY gained, still within the range for it to be considered "cost-effective" (i.e., under $50,000-$100,000/DQALY gained).

Table 9. Procedure Related Costs

Procedure	1998 $	Published range	
		Low	High
Fine needle aspiration	212	75	320
Core needle biopsy	286	NA	
Excision of breast lesion[6]	1,272	702	1410
Mammogram directed	349	NA	NA
Ultrasound directed	386	NA	NA

[6] Includes $500 facility cost

Table 10. Results of Cost-effectiveness Analysis

Strategy	Discounted lifetime costs (3%/yr)	Discounted quality-adjusted life expectancy (DQALY*)	Marginal cost-effectiveness ratio ($/DQALY gained)
FNA positive if malignant or suspicious cytology	27,224	17.34	
FNA positive if malignant, suspicious or atypical cytology	27,235	17.35	760
FNA positive only if malignant cytology	27,381	17.27	Inferior
Open biopsy	27,471	17.37	11,900

*DQALY = discounted quality-adjusted life year

Inferior = Higher cost and lower life expectancy than next more costly strategy

Key Research Accomplishments

I. The performance of a comprehensive and structured literature review of fine needle aspiration for breast lesions, the largest meta-analysis performed in this area to date during years one and two of the project.

II. The conduct of physician focus groups in years one and two of the project.

III. Construction of a decision analysis that compares fine-needle aspiration compared to open biopsy taking into consideration false negative, false positive cytological results and the long-term clinical outcomes in year 2.

IV. Comparison of the cost-effectiveness of fine-needle aspiration compared to open biopsy taking into consideration both the short- and long-term economic and clinical outcomes in year two.

Reportable Outcomes

We have no publications to date but anticipate submitting 4 manuscripts based on this work to report separately the results of the meta-analysis for mammography directed fine needle aspiration, ultrasound directed fine needle aspiration and results for those that were not assisted by an imaging modality. Based on the meta-analysis, we will submit a manuscript comparing the cost-effectiveness of fine needle aspiration to open biopsy.

The database of articles retrieved will provide a rich source for potential future analyses, such as a comparative meta-analysis which include studies which directly examine alternative methodologies for diagnosing breast lesions.

Conclusions

Our results constitute the largest meta-analysis performed in this area to date. It suggests substantial variation in sensitivity and specificity in the performance of fine needle aspirate for evaluation of breast lesions with the estimates being lower than that reported in some prior reports.[103] Sources for variation in its test characteristics include patient differences and interpretation differences among cytopathologists. This study, however, also demonstrates study to study variation in equipment and technique including the size syringe and the gauge needle used, the fixative and the stain applied and whether centrifugation of fluid was performed. Despite these differences, no randomized controlled studies have been performed to compare these various methodologic techniques for their sensitivity and specificity.[104, 105] Our results emphasize the importance for the local facilities to determine their sensitivity and specificity in a series of unselected patients confirmed by biopsy to estimate the local experience and to determine if aspiration cytology is appropriate.

Our results define cytological results as positive or negative. Alternatively, we could have calculated likelihood ratios for each category, i.e., the likelihood of any specific cytological interpretation among patients with histologically defined malignancy compared to those with benign histology.[106] Because not all studies use the same categories for reporting cytology results, however, such an analysis would result in biased estimates for the less common categories, such as suspicious and atypical. Instead, our methodology for analysis allows the incorporation of more studies for each estimate of sensitivity and specificity. Moreover, as would occur clinically, our analysis suggests that positive results would increase the likelihood of a malignant etiology underlying the breast lesion and would necessitate more definitive evaluation such as a breast biopsy.

This analysis is limited by the absence of gold standard testing in all patients. Patients included in this analysis all had biopsies performed, most likely influenced by patient and clinical characteristics that led their physicians to seek definitive histologic confirmation. The absence of complete histologic confirmation in all cases is termed verification bias. Although methodologies exist to attempt to adjust for such bias, such corrections are likely to be biased because they assume that selection for the second test is unbiased. Clinicians select patients for histologic confirmation based on their risk factors for breast cancer, e.g., family history or lesion characteristics by palpation or imaging. Any correction for verification bias then most likely overcorrects. The true sensitivity and specificity likely lies between the unadjusted and the adjusted for verification estimates.

The meta-analysis does not answer whether a fine needle aspirate or a biopsy should be done. Such a comparison could be addressed by a randomized trial comparing the two approaches, but even so, there may be local variation (procedure performance or pathologic interpretation) that may influence the optimal procedure in a particular locale. To assist with such determination, our study continues in a second phase to examine the relative costs and benefits of fine needle aspiration in terms of life expectancy and lifetime costs. In particular, false positive results lead to morbidity from anxiety and economic costs because of unnecessary biopsy. Those with false negative cytological

results have a delay in the accurate diagnosis of their breast cancer, potentially leading to a lower life expectancy and increased cost of care when they present with more advanced breast cancer. Thus, this serves as the rationale for the second half of our grant to examine the cost-effectiveness of fine needle aspiration compared to open biopsy.

Our results are consistent with expert opinion regarding the cutoff value that should be used to pursue biopsy. Consistent with our physician focus groups, patient with atypical as well as malignant or suspicious cytology results should have breast biopsy pursued because of the risk for false negative results (by excluding these patients). The long-term economic and clinical effects outweigh the short-term risks. Women's feelings about open biopsy versus fine needle aspiration may also influence the choice and deserves further study.

Nearly all previous analyses have simply examined procedure related costs, comparing cost savings from fine needle aspiration to open biopsy. Our analysis considers also the potential delay in the diagnosis for those who have false negative results with fine needle aspiration. The effect of this delay in diagnostic staging of the disease is not known. Based on the assumptions of our model, if half of the patients who have a delay exceeding 3 months advance from local to regional disease because of the false negative cytology, then open biopsy might be preferred over fine needle aspiration.

On the other hand, in a young patient with a breast cyst and a low likelihood of cancer, our results support the use of fine needle aspiration because in such situations, the cost-effectiveness of breast biopsy exceeds $50,000 per discounted quality-adjusted life year gained. Our results suggest that future studies examining the effect of false negative cytology results on the stage of breast cancer at the time of delayed presentation would be an important factor in deciding whether to opt for fine needle aspiration or open biopsy. Lastly, the analysis of tru-cut biopsy is beyond the scope of this study, but a comparison study should be undertaken given the small differences in cost compared to fine needle aspiration and presumed higher sensitivity and specificity.

References

1. Roe SM, Mathews JA, Burns RP, Sumida MP, Craft P Jr, Greer MS. Stereotactic and ultrasound core needle breast biopsy performed by surgeons. American Journal of Surgery 1997; 174:699-703; discussion 703-4.
2. Abele JS, Miller TR, Goodson WH, 3d, Hunt TK, Hohn DC. Fine-needle aspiration of palpable breast masses. A program for staged implementation. Archives of Surgery 1983; 118:859-63.
3. Adye B, Jolly PC, Bauermeister DE. The role of fine-needle aspiration in the management of solid breast masses. Archives of Surgery 1988; 123:37-9.
4. Aretz HT, Silverman ML, Kolodziejski JL, Witherspoon BR. Fine-needle aspiration. Why it deserves another look. Postgraduate Medicine 1984; 75:49-52, 55-6.
5. Arishita GI, Cruz BK, Harding CT, Arbutina DR. Mammogram-directed fine-needle aspiration of nonpalpable breast lesions. Journal of Surgical Onçology 1991; 48:153-7.
6. Atamdede FI, Isaacs JH. The role of fine-needle aspiration in the diagnosis of breast lesions. Gynecologic Oncology 1993; 50:159-63.
7. Barrows GH, Anderson TJ, Lamb JL, Dixon JM. Fine-needle aspiration of breast cancer. Relationship of clinical factors to cytology results in 689 primary malignancies. Cancer 1986; 58:1493-8.
8. Bibbo M, Zuspan FP. Fine-needle aspiration of the breast in an obstetrics and gynecology hospital. American Journal of Obstetrics & Gynecology 1975; 122:525-6.
9. Bodo M, Bak M. Aspiration cytodiagnostics of breast tumours. Archiv fur Geschwulstforschung 1984; 54:219-24.
10. Bucchi L, Schincaglia P, Melandri G, et al. Performance of fineneedle aspiration cytology of the breast. Clinical experience in Ravenna (Italy). Tumori 1993; 79:413-7.
11. Cappellari A, Bagarella M, Corradi G. Application of monoclonal antibodies in diagnosis of human breast cancer by fine-needle aspiration cytology. Oncology 1990; 47:398-400.
12. Casey TT, Rodgers WH, Baxter JW, Sawyers JL, Reynolds VH, Page DL. Stratified diagnostic approach to fine needle aspiration of the breast. American Journal of Surgery 1992; 163:305-11.
13. Cheung PS, Yan KW, Alagaratnam TT. The complementary role of fine needle aspiration cytology and Tru-cut needle biopsy in the management of breast masses. Australian & New Zealand Journal of Surgery 1987; 57:615-20.
14. Ciatto S, Rosselli Del Turco M, Bravetti P. Nonpalpable breast lesions: stereotaxic fine-needle aspiration cytology. Radiology 1989; 173:57-9.
15. Cook AI, Robinson KB. Aspiration cytology of the breast with immediate reporting: initial experience with 1000 patients in a district general hospital. Journal of the Royal College of Surgeons of Edinburgh 1991; 36:289-92.
16. Costa MJ, Tadros T, Hilton G, Birdsong G. Breast fine needle aspiration cytology. Utility as a screening tool for clinically palpable lesions. Acta Cytologica 1993; 37:461-71.

17. Crone P, Hertz J, Nilsson T, et al. The predictive value of three diagnostic procedures in the evaluation of palpable breast tumours. Annales Chirurgiae et Gynaecologiae 1984; 73:273-6.
18. Daidone MG, Martelli G, Pilotti S, et al. Contribution of ploidy and cell kinetics from fine-needle aspirates for the diagnosis of breast lesions: study of 606 consecutive cases. Cytometry 1995; 22:177-80.
19. Daum GS, Kline TS, Artymyshyn RL, Neal HS. Aspiration biopsy cytology of occult breast lesions by use of the "scouting needle". A prospective study of 261 cases. Cancer 1991; 67:2150-2.
20. Dehn TC, Clarke J, Dixon JM, Crucioli V, Greenall MJ, Lee EC. Fine needle aspiration cytology, with immediate reporting, in the outpatient diagnosis of breast disease. Annals of the Royal College of Surgeons of England 1987; 69:280-2.
21. Deschenes L, Fabia J, Meisels A, et al. Fine needle aspiration biopsy in the management of palpable breast lesions. Canadian Journal of Surgery 1978; 21:417-9.
22. Dowlatshahi K, Yaremko ML, Kluskens LF, Jokich PM. Nonpalpable breast lesions: findings of stereotaxic needle-core biopsy and fine-needle aspiration cytology. Radiology 1991; 181:745-50.
23. Elfrink RJ, Ulbrich J, Metzler MH. Fine-needle aspiration of breast masses. Missouri Medicine 1994; 91:293-8.
24. Elston CW, Cotton RE, Davies CJ, Blamey RW. A comparison of the use of the "Tru-Cut" needle and fine needle aspiration cytology in the pre-operative diagnosis of carcinoma of the breast. Histopathology 1978; 2:239-54.
25. Fajardo LL, Davis JR, Wiens JL, Trego DC. Mammography-guided stereotactic fine-needle aspiration cytology of nonpalpable breast lesions: prospective comparison with surgical biopsy results. AJR 1990; 155:977-81.
26. Fessia L, Botta G, Arisio R, Verga M, Aimone V. Fine-needle aspiration of breast lesions: role and accuracy in a review of 7,495 cases. Diagnostic Cytopathology 1987; 3:121-5.
27. Fornage BD, Faroux MJ, Simatos A. Breast masses: US-guided fine-needle aspiration biopsy. Radiology 1987; 162:409-14.
28. Gardecki TI, Hogbin BM, Melcher DH, Smith RS. Aspiration cytology in the preoperative management of breast cancer. Lancet 1980; 2:790-2.
29. Gelabert HA, Hsiu JG, Mullen JT, Jaffe AH, D'Amato NA. Prospective evaluation of the role of fine-needle aspiration biopsy in the diagnosis and management of patients with palpable solid breast lesions. American Surgeon 1990; 56:263-7.
30. Gent HJ, Denkert B, Lagemann A. Stereotaxic fine needle aspiration of nonpalpable breast lesions with the cytoguide unit: first experiences. Recent Results in Cancer Research 1990; 119:109-13.
31. Gordon PB, Goldenberg SL, Chan NH. Solid breast lesions: diagnosis with US-guided fine-needle aspiration biopsy. Radiology 1993; 189:573-80.
32. Griffith CN, Kern WH, Mikkelsen WP. Needle aspiration cytologic examination in the management of suspicious lesions of the breast. Surgery, Gynecology & Obstetrics 1986; 162:142-4.

33. Hann L, Ducatman BS, Wang HH, Fein V, McIntire JM. Nonpalpable breast lesions: evaluation by means of fine-needle aspiration cytology. Radiology 1989; 171:373-6.
34. Hatada T, Aoki I, Okada K, Nakai T, Utsunomiya J. Usefulness of ultrasound-guided, fine-needle aspiration biopsy for palpable breast tumors. Archives of Surgery 1996; 131:1095-8.
35. Horgan PG, Waldron D, Mooney E, O'Brien D, McGuire M, Given HF. The role of aspiration cytologic examination in the diagnosis of carcinoma of the breast. Surgery, Gynecology & Obstetrics 1991; 172:290-2.
36. Iles SE, MacGregor JH, Bodurtha AJ, Bernardo AI, Daniels C. Stereotactic fine-needle aspiration cytology of nonpalpable breast lesions: initial experience in a tertiary-care institution. Canadian Association of Radiologists Journal 1994; 45:28-34.
37. Kahky MP, Rone VR, Duncan DL, Cruz AB, Jr, Gaskill HV, 3d. Needle aspiration biopsy of palpable breast masses. American Journal of Surgery 1988; 156:450-2.
38. Kataoka T, Nishiki M, Matsuyama T, Yamane M, Kishi N, Dohi K. Diagnosis of breast lesions by aspiration biopsy cytology. Japanese Journal of Surgery 1985; 15:361-7.
39. Kerin MJ, McAnena OJ, Waldron RP, McGuire M, Given HF. Diagnostic pitfalls of fine needle aspiration cytology for breast disease. Irish Medical Journal 1993; 86:100-1.
40. Kher AV, Marwar AW, Raichur BS. "Evaluation of fine needle aspiration biopsy in the diagnosis of breast lesions". Indian Journal of Pathology & Microbiology 1981; 24:100-106c.
41. Knight DC, Lowell DM, Heimann A, Dunn E. Aspiration of the breast and nipple discharge cytology. Surgery, Gynecology & Obstetrics 1986; 163:415-20.
42. Koivuniemi AP. Fine-needle aspiration biopsy of the breast. Annals of Clinical Research 1976; 8:272-83.
43. Kreuzer G, Boquoi E. Aspiration biopsy cytology, mammography and clinical exploration: a modern set up in diagnosis of tumors of the breast. Acta Cytologica 1976; 20:319-23.
44. Kwok D, Chan M, Gwi E, Law D. Aspiration cytology in the management of breast lesions. Australian & New Zealand Journal of Surgery 1988; 58:295-9.
45. Learmonth GM, Hayes MM, Hacking A, et al. Fine-needle aspiration biopsy cytology of the breast. A review of the Groote Schuur Hospital experience. South African Medical Journal 1987; 72:525-7.
46. Lew WY, Lee WH. Fine needle aspiration cytology: its role in the management of breast tumours. Australian & New Zealand Journal of Surgery 1988; 58:941-6.
47. Lioe TF, Elliott H, Allen DC, Spence RA. A 3 year audit of fine needle aspirates from a symptomatic breast clinic. Ulster Medical Journal 1997; 66:24-7.
48. Malberger E, Toledano C, Barzilai A, Schramek A. The decisive role of fine needle aspiration cytology in the preoperative work-up of breast cancer. Israel Journal of Medical Sciences 1981; 17:899-904.
49. Marasa L, Tomasino RM. Aspiration cytology of the breast. I. Comparison of cytologic and histologic findings. Pathologica 1982; 74:183-92.

50. Marcaccio MJ, O'Brien SE, Chen VS. Fine-needle aspiration cytology in breast lumps. Canadian Journal of Surgery 1986; 29:405-7.
51. Masood S, Frykberg ER, McLellan GL, Scalapino MC, Mitchum DG, Bullard JB. Prospective evaluation of radiologically directed fine-needle aspiration biopsy of nonpalpable breast lesions. Cancer 1990; 66:1480-7.
52. McDonald AH, Hanchard B, Shah D, Fletcher PR, DuQuesnay R. Fine needle aspiration biopsy in the diagnosis of breast cancer. West Indian Medical Journal 1990; 39:71-3.
53. Miccoli P, Iacconi P, Simi U, Roncella M, Arganini M, Melis GB. Fine needle aspiration in the diagnosis of breast masses: our experience. European Journal of Surgical Oncology 1986; 12:123-6.
54. Mitnick JS, Vazquez MF, Roses DF, Harris MN, Gianutsos R, Waisman J. Stereotaxic localization for fine-needle aspiration breast biopsy. Initial experience with 300 patients. Archives of Surgery 1991; 126:1137-40.
55. Nicastri GR, Reed WP, Dziura BR. The accuracy of malignant diagnoses established by fine needle aspiration cytologic procedures of mammary masses. Surgery, Gynecology & Obstetrics 1991; 172:457-60.
56. Nicosia SV, Williams JA, Horowitz SA, et al. Fine needle aspiration biopsy of palpable breast lesions. Review and statistical analysis of 1875 cases. Surgical Oncology 1993; 2:145-60.
57. Ogunniyi JO, Senbanjo RO, Ogunlusi ML. Fine needle aspiration cytology in the assessment of breast lumps in Ibadan. African Journal of Medicine & Medical Sciences 1989; 18:151-4.
58. Owen AW, Anderson TJ, Forrest AP. "Closed" biopsy for breast cancer. Journal of the Royal College of Surgeons of Edinburgh 1980; 25:237-41.
59. Place R, Velanovich V, Carter P. Fine needle aspiration in the clinical management of mammary masses. Surgery, Gynecology & Obstetrics 1993; 177:7-11.
60. Rotstein S, Nilsson B, Svane G, Gustavson-Kadaka E. Clinical examination, mammographic findings and cytological diagnosis in patients with breast disorders. Results of 9 years' follow-up. Acta Oncologica 1992; 31:393-7.
61. Saarela AO, Kiviniemi HO, Rissanen TJ, Paloneva TK. Nonpalpable breast lesions: pathologic correlation of ultrasonographically guided fine-needle aspiration biopsy. Journal of Ultrasound in Medicine 1996; 15:549-53; quiz 555-6.
62. Saunders G, Lakra Y, Libcke J. Comparison of needle aspiration cytologic diagnosis with excisional biopsy tissue diagnosis of palpable tumors of the breast in a community hospital. Surgery, Gynecology & Obstetrics 1991; 172:437-40.
63. Scopa CD, Koukouras D, Spiliotis J, et al. Comparison of fine needle aspiration and Tru-Cut biopsy of palpable mammary lesions. Cancer Detection & Prevention 1996; 20:620-4.
64. Sharifah NA, Hayati AR, Meah F, al-Rashid Z. Fine needle aspiration cytology of the breast: an analysis of 1,094 aspirates. Malaysian Journal of Pathology 1993; 15:53-8.

65. Sheridan WG, Davies JE, Lewis MH. Early experience with fine-needle aspiration cytology in a district general hospital. British Journal of Hospital Medicine 1987; 37:156-8.
66. Sneige N, Holder PD, Katz RL, et al. Fine-needle aspiration cytology of the male breast in a cancer center. Diagnostic Cytopathology 1993; 9:691-7.
67. Somers RG, Young GP, Kaplan MJ, Bernhard VM, Rosenberg M, Somers D. Fine-needle aspiration biopsy in the management of solid breast tumors. Archives of Surgery 1985; 120:673-7.
68. Stavric GD, Tevcev DT, Kaftandjiev DR, Novak JJ. Aspiration biopsy cytologic method in diagnosis of breast lesions: a critical review of 250 cases. Acta Cytologica 1973; 17:188-90.
69. Steinberg JL, Trudeau ME, Ryder DE, et al. Combined fine-needle aspiration, physical examination and mammography in the diagnosis of palpable breast masses: their relation to outcome for women with primary breast cancer. Canadian Journal of Surgery 1996; 39:302-11.
70. Strawbridge HT, Bassett AA, Foldes I. Role of cytology in management of lesions of the breast. Surgery, Gynecology & Obstetrics 1981; 152:1-7.
71. Svane G. Stereotaxic fine needle biopsy of nonpalpable breast lesions performed by the Mammotest. Recent Results in Cancer Research 1987; 105:95-6.
72. Teixidor HS, Wojtasek DA, Reiches EM, Santos-Buch CA, Minick CR. Fine-needle aspiration of breast biopsy specimens: correlation of histologic and cytologic findings. Radiology 1992; 184:55-8.
73. Thomas PA, Vazquez MF, Waisman J. Comparison of fine-needle aspiration and frozen section of palpable mammary lesions. Modern Pathology 1990; 3:570-4.
74. Ulanow RM, Galblum L, Canter JW. Fine needle aspiration in the diagnosis and management of solid breast lesions. American Journal of Surgery 1984; 148:653-7.
75. Verma K, Kapila K. The role of fine needle aspiration cytology of breast lumps in the management of patients. Indian Journal of Medical Research 1989; 90:135-9.
76. Vetrani A, Fulciniti F, Di Benedetto G, et al. Fine-needle aspiration biopsies of breast masses. An additional experience with 1153 cases (1985 to 1988) and a meta-analysis. Cancer 1992; 69:736-40.
77. Wallgren A, Zajicek J. Cytologic presentation of mammary carcinoma on aspiration biopsy smears. Acta Cytologica 1976; 20:469-78.
78. Wilson SL, Ehrmann RL. The cytologic diagnosis of breast aspirations. Acta Cytologica 1978; 22:470-5.
79. Wollenberg NJ, Caya JG, Clowry LJ. Fine needle aspiration cytology of the breast. A review of 321 cases with statistical evaluation. Acta Cytologica 1985; 29:425-9.
80. Young GP, Somers RG, Young I, Kaplan M, Cowan DF. Experience with a modified fine-needle aspiration biopsy technique in 533 breast cases. Diagnostic Cytopathology 1986; 2:91-8.
81. Yu SC, Lin PW, Liaw KY, et al. Fine needle aspiration cytology of breast tumors. Taiwan I Hsueh Hui Tsa Chih - Journal of the Formosan Medical Ass 1985; 84:727-34.

82. Zajdela A, Zillhardt P, Voillemot N. Cytological diagnosis by fine needle sampling without aspiration. Cancer 1987; 59:1201-5.
83. Zemba-Palko V, Minkowitz G, Kline TS. Occult breast lesions and aspiration biopsy: the Lankenau experience. Diagnostic Cytopathology 1994; 11:15-9.
84. Beck JR, Pauker SG, Gottlieb JE, Klein K, Kassirer JP. A convenient approximation of life expectancy (the DEALE): II. Use in medical decision making. American Journal of Medicine 1982; 73:889-897.
85. Beck JR, Pauker SG. The Markov Process in medical prognosis. Medical Decision Making 1983; 3:419-58.
86. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Medical Decision Making 1993; 13:322-338.
87. Goodson WH, 3d, Mailman R, Miller TR. Three year follow-up of benign fine-needle aspiration biopsies of the breast. American Journal of Surgery 1987; 154:58-61.
88. Bates AT, Bates T, Hastrich DJ, et al. Delay in the diagnosis of breast cancer: the effect of the introduction of fine needle aspiration cytology to a breast clinic. European Journal of Surgical Oncology 1992; 18:433-7.
89. De Haes JCJM, De Koning HJ, van Oortmarssen GJ, Van Agt HME, De Bruyn AE, Van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life years. International Journal of Cancer 1991; 49:538-544.
90. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. Journal of Chronic Disease 1978; 31:697-704.
91. Wasserman Y. Physicians' Fee Reference 1998. West Allis, WI: Medical Publishers, Ltd, 1998.
92. Hartunian NS, Smart CN, Thompson MS. The incidence and economic costs of cancer, motor vehicle injuries, coronary heart disease, and stroke: a comparative analysis. American Journal of Public Health 1980; 70:1249-60.
93. Rice DP. Estimating the costs of illness. HS Publ No. 947-6. Washington, DC: Home Economics Branch, Division of Medical Care Administration, PHS, DEW, 1966.
94. Fireman BH, Quesenberry CP, CP Somkin, et al. Cost of care for cancer in a health maintenance organization. Health Care Financing Reviews 1997; 18:51-76.
95. Baker MS, Kessler LG, Urban N, Smucker RC. Estimating the treatment costs of breast and lung cancer. Medical Care 1991; 29:40-9.
96. Eddy DM, Hasselblad V, McGivney W, Hendee W. The value of mammography screening in women under age 50 years. JAMA 1988; 259:1512-9.
97. Taplin SH, Barlow W, Urban N, et al. Stage, age, comorbidity, and direct costs of colon, prostate, and breast cancer care. Journal of the National Cancer Institute 1995; 87:417-26.
98. Smith TJ, Safaii H, Foster EA, Reinhold RB. Accuracy and cost-effectiveness of fine needle aspiration biopsy. American Journal of Surgery 1985; 149:540-5.
99. Silverman JF, Lannin DR, O'Brien K, Norris HT. The triage role of fine needle aspiration biopsy of palpable breast masses. Diagnostic accuracy and cost-effectiveness. Acta Cytologica 1987; 31:731-6.
100. Velanovich V. Fine needle aspiration cytology in the diagnosis and management of ectopic breast tissue. American Surgeon 1995; 61:277-8.

101. Lannin DR, Silverman JF, Walker C, Pories WJ. Cost-effectiveness of fine needle biopsy of the breast. Annals of Surgery 1986; 203:474-80.
102. Logan-Young W, Dawson AE, Wilbur DC, et al. The cost-effectiveness of fine-needle aspiration cytology and 14-gauge core needle biopsy compared with open surgical biopsy in the diagnosis of breast carcinoma. Cancer 1998; 82:1867-73.
103. Mushlin AI. Diagnostic tests in breast cancer. Clinical Strategies based on diagnostic probabilities. Annals of Internal Medicine 1985; 103:79-85.
104. Zarbo RJ, Howanitz PJ, Bachner P. Interinstitutional comparison of performance in breast fine-needle aspiration cytology. A Q-probe quality indicator study. Archives of Pathology & Laboratory Medicine 1991; 115:743-50.
105. Anonymous. The uniform approach to breast fine-needle aspiration biopsy. NIH Consensus Development Conference. [91 refs]. American Journal of Surgery 1997; 174:371-85.
106. Giard RW, Hermans J. Interpretation of diagnostic cytology with likelihood ratios. Archives of Pathology & Laboratory Medicine 1990; 114:852-4.

Appendix I: Meta-analysis Review Form

Article #:
First author (last name, first initial):
Institution:
Reviewer:

Rejection criteria:

1) duplicate publication
2) N<50
3) Absence of primary data
4) Special population
5) Absent reference or gold standard
6) Inadequate detail to allow determination of sensitivity or specificity

United States:
Department:

- Family Practice
- Gynecology
- Oncology
- Pathology
- Radiation oncology
- Radiology
- Surgery
- Other
- Not specified

Years study done

- Initial
- Final

Lesion characteristics

- Solid specified

Localization:

- Ultrasound
- Mammography
- Stereotaxic
- Other

Technique:

- Gauge needle specified (list all):
- Millimeter syringe specified (list all):
- Number of aspirations from lesion specified:
- Fixitive:
  - Air dried
  - Fixitive
  - None specified
- Fluid centrifuged
- Stain

Pap
Diff Quik
Giemsa
Hematoxylin and eosin
Romanovsky
Wright
Other

Patient characteristics
# of patients
# of male patients
Youngest
Oldest
Mean or median

Fine needle aspirations
# of aspirations

Lesion biopsies
# of biopsies

All aspirations biopsied

Results Complete 2 x 5 table:

Biopsy Results	Cytology Results				
	Malignant	Suspicous	Atypical	Benign	Inadequate
Malignant					
Benign					

Follow-up duration (months):

Appendix II. Studies Included in Meta-analysis

1. Abele JS, Miller TR, Goodson WH, 3d, Hunt TK, Hohn DC. Fine-needle aspiration of palpable breast masses. A program for staged implementation. Archives of Surgery 1983; 118:859-63.
2. Adye B, Jolly PC, Bauermeister DE. The role of fine-needle aspiration in the management of solid breast masses. Archives of Surgery 1988; 123:37-9.
3. Aretz HT, Silverman ML, Kolodziejski JL, Witherspoon BR. Fine-needle aspiration. Why it deserves another look. Postgraduate Medicine 1984; 75:49-52, 55-6.
4. Arishita GI, Cruz BK, Harding CT, Arbutina DR. Mammogram-directed fine-needle aspiration of nonpalpable breast lesions. Journal of Surgical Oncology 1991; 48:153-7.
5. Atamdede FI, Isaacs JH. The role of fine-needle aspiration in the diagnosis of breast lesions. Gynecologic Oncology 1993; 50:159-63.
6. Barrows GH, Anderson TJ, Lamb JL, Dixon JM. Fine-needle aspiration of breast cancer. Relationship of clinical factors to cytology results in 689 primary malignancies. Cancer 1986; 58:1493-8.
7. Bibbo M, Zuspan FP. Fine-needle aspiration of the breast in an obstetrics and gynecology hospital. American Journal of Obstetrics & Gynecology 1975; 122:525-6.
8. Bodo M, Bak M. Aspiration cytodiagnostics of breast tumours. Archiv fur Geschwulstforschung 1984; 54:219-24.
9. Bucchi L, Schincaglia P, Melandri G, et al. Performance of fineneedle aspiration cytology of the breast. Clinical experience in Ravenna (Italy). Tumori 1993; 79:413-7.
10. Cappellari A, Bagarella M, Corradi G. Application of monoclonal antibodies in diagnosis of human breast cancer by fine-needle aspiration cytology. Oncology 1990; 47:398-400.
11. Casey TT, Rodgers WH, Baxter JW, Sawyers JL, Reynolds VH, Page DL. Stratified diagnostic approach to fine needle aspiration of the breast. American Journal of Surgery 1992; 163:305-11.
12. Cheung PS, Yan KW, Alagaratnam TT. The complementary role of fine needle aspiration cytology and Tru-cut needle biopsy in the management of breast masses. Australian & New Zealand Journal of Surgery 1987; 57:615-20.
13. Ciatto S, Rosselli Del Turco M, Bravetti P. Nonpalpable breast lesions: stereotaxic fine-needle aspiration cytology. Radiology 1989; 173:57-9.
14. Cook AI, Robinson KB. Aspiration cytology of the breast with immediate reporting: initial experience with 1000 patients in a district general hospital. Journal of the Royal College of Surgeons of Edinburgh 1991; 36:289-92.
15. Costa MJ, Tadros T, Hilton G, Birdsong G. Breast fine needle aspiration cytology. Utility as a screening tool for clinically palpable lesions. Acta Cytologica 1993; 37:461-71.
16. Crone P, Hertz J, Nilsson T, et al. The predictive value of three diagnostic procedures in the evaluation of palpable breast tumours. Annales Chirurgiae et Gynaecologiae 1984; 73:273-6.

17. Daidone MG, Martelli G, Pilotti S, et al. Contribution of ploidy and cell kinetics from fine-needle aspirates for the diagnosis of breast lesions: study of 606 consecutive cases. Cytometry 1995; 22:177-80.
18. Daum GS, Kline TS, Artymyshyn RL, Neal HS. Aspiration biopsy cytology of occult breast lesions by use of the "scouting needle". A prospective study of 261 cases. Cancer 1991; 67:2150-2.
19. Dehn TC, Clarke J, Dixon JM, Crucioli V, Greenall MJ, Lee EC. Fine needle aspiration cytology, with immediate reporting, in the outpatient diagnosis of breast disease. Annals of the Royal College of Surgeons of England 1987; 69:280-2.
20. Deschenes L, Fabia J, Meisels A, et al. Fine needle aspiration biopsy in the management of palpable breast lesions. Canadian Journal of Surgery 1978; 21:417-9.
21. Dowlatshahi K, Yaremko ML, Kluskens LF, Jokich PM. Nonpalpable breast lesions: findings of stereotaxic needle-core biopsy and fine-needle aspiration cytology. Radiology 1991; 181:745-50.
22. Elfrink RJ, Ulbrich J, Metzler MH. Fine-needle aspiration of breast masses. Missouri Medicine 1994; 91:293-8.
23. Elston CW, Cotton RE, Davies CJ, Blamey RW. A comparison of the use of the "Tru-Cut" needle and fine needle aspiration cytology in the pre-operative diagnosis of carcinoma of the breast. Histopathology 1978; 2:239-54.
24. Fajardo LL, Davis JR, Wiens JL, Trego DC. Mammography-guided stereotactic fine-needle aspiration cytology of nonpalpable breast lesions: prospective comparison with surgical biopsy results. AJR 1990; 155:977-81.
25. Fessia L, Botta G, Arisio R, Verga M, Aimone V. Fine-needle aspiration of breast lesions: role and accuracy in a review of 7,495 cases. Diagnostic Cytopathology 1987; 3:121-5.
26. Fornage BD, Faroux MJ, Simatos A. Breast masses: US-guided fine-needle aspiration biopsy. Radiology 1987; 162:409-14.
27. Gardecki TI, Hogbin BM, Melcher DH, Smith RS. Aspiration cytology in the preoperative management of breast cancer. Lancet 1980; 2:790-2.
28. Gelabert HA, Hsiu JG, Mullen JT, Jaffe AH, D'Amato NA. Prospective evaluation of the role of fine-needle aspiration biopsy in the diagnosis and management of patients with palpable solid breast lesions. American Surgeon 1990; 56:263-7.
29. Gent HJ, Denkert B, Lagemann A. Stereotaxic fine needle aspiration of nonpalpable breast lesions with the cytoguide unit: first experiences. Recent Results in Cancer Research 1990; 119:109-13.
30. Gordon PB, Goldenberg SL, Chan NH. Solid breast lesions: diagnosis with US-guided fine-needle aspiration biopsy. Radiology 1993; 189:573-80.
31. Griffith CN, Kern WH, Mikkelsen WP. Needle aspiration cytologic examination in the management of suspicious lesions of the breast. Surgery, Gynecology & Obstetrics 1986; 162:142-4.
32. Hann L, Ducatman BS, Wang HH, Fein V, McIntire JM. Nonpalpable breast lesions: evaluation by means of fine-needle aspiration cytology. Radiology 1989; 171:373-6.

33. Hatada T, Aoki I, Okada K, Nakai T, Utsunomiya J. Usefulness of ultrasound-guided, fine-needle aspiration biopsy for palpable breast tumors. Archives of Surgery 1996; 131:1095-8.
34. Horgan PG, Waldron D, Mooney E, O'Brien D, McGuire M, Given HF. The role of aspiration cytologic examination in the diagnosis of carcinoma of the breast. Surgery, Gynecology & Obstetrics 1991; 172:290-2.
35. Iles SE, MacGregor JH, Bodurtha AJ, Bernardo AI, Daniels C. Stereotactic fine-needle aspiration cytology of nonpalpable breast lesions: initial experience in a tertiary-care institution. Canadian Association of Radiologists Journal 1994; 45:28-34.
36. Kahky MP, Rone VR, Duncan DL, Cruz AB, Jr, Gaskill HV, 3d. Needle aspiration biopsy of palpable breast masses. American Journal of Surgery 1988; 156:450-2.
37. Kataoka T, Nishiki M, Matsuyama T, Yamane M, Kishi N, Dohi K. Diagnosis of breast lesions by aspiration biopsy cytology. Japanese Journal of Surgery 1985; 15:361-7.
38. Kerin MJ, McAnena OJ, Waldron RP, McGuire M, Given HF. Diagnostic pitfalls of fine needle aspiration cytology for breast disease. Irish Medical Journal 1993; 86:100-1.
39. Kher AV, Marwar AW, Raichur BS. "Evaluation of fine needle aspiration biopsy in the diagnosis of breast lesions". Indian Journal of Pathology & Microbiology 1981; 24:100-106c.
40. Knight DC, Lowell DM, Heimann A, Dunn E. Aspiration of the breast and nipple discharge cytology. Surgery, Gynecology & Obstetrics 1986; 163:415-20.
41. Koivuniemi AP. Fine-needle aspiration biopsy of the breast. Annals of Clinical Research 1976; 8:272-83.
42. Kreuzer G, Boquoi E. Aspiration biopsy cytology, mammography and clinical exploration: a modern set up in diagnosis of tumors of the breast. Acta Cytologica 1976; 20:319-23.
43. Kwok D, Chan M, Gwi E, Law D. Aspiration cytology in the management of breast lesions. Australian & New Zealand Journal of Surgery 1988; 58:295-9.
44. Learmonth GM, Hayes MM, Hacking A, et al. Fine-needle aspiration biopsy cytology of the breast. A review of the Groote Schuur Hospital experience. South African Medical Journal 1987; 72:525-7.
45. Lew WY, Lee WH. Fine needle aspiration cytology: its role in the management of breast tumours. Australian & New Zealand Journal of Surgery 1988; 58:941-6.
46. Lioe TF, Elliott H, Allen DC, Spence RA. A 3 year audit of fine needle aspirates from a symptomatic breast clinic. Ulster Medical Journal 1997; 66:24-7.
47. Malberger E, Toledano C, Barzilai A, Schramek A. The decisive role of fine needle aspiration cytology in the preoperative work-up of breast cancer. Israel Journal of Medical Sciences 1981; 17:899-904.
48. Marasa L, Tomasino RM. Aspiration cytology of the breast. I. Comparison of cytologic and histologic findings. Pathologica 1982; 74:183-92.
49. Marcaccio MJ, O'Brien SE, Chen VS. Fine-needle aspiration cytology in breast lumps. Canadian Journal of Surgery 1986; 29:405-7.

50. Masood S, Frykberg ER, McLellan GL, Scalapino MC, Mitchum DG, Bullard JB. Prospective evaluation of radiologically directed fine-needle aspiration biopsy of nonpalpable breast lesions. Cancer 1990; 66:1480-7.
51. McDonald AH, Hanchard B, Shah D, Fletcher PR, DuQuesnay R. Fine needle aspiration biopsy in the diagnosis of breast cancer. West Indian Medical Journal 1990; 39:71-3.
52. Miccoli P, Iacconi P, Simi U, Roncella M, Arganini M, Melis GB. Fine needle aspiration in the diagnosis of breast masses: our experience. European Journal of Surgical Oncology 1986; 12:123-6.
53. Mitnick JS, Vazquez MF, Roses DF, Harris MN, Gianutsos R, Waisman J. Stereotaxic localization for fine-needle aspiration breast biopsy. Initial experience with 300 patients. Archives of Surgery 1991; 126:1137-40.
54. Nicastri GR, Reed WP, Dziura BR. The accuracy of malignant diagnoses established by fine needle aspiration cytologic procedures of mammary masses. Surgery, Gynecology & Obstetrics 1991; 172:457-60.
55. Nicosia SV, Williams JA, Horowitz SA, et al. Fine needle aspiration biopsy of palpable breast lesions. Review and statistical analysis of 1875 cases. Surgical Oncology 1993; 2:145-60.
56. Ogunniyi JO, Senbanjo RO, Ogunlusi ML. Fine needle aspiration cytology in the assessment of breast lumps in Ibadan. African Journal of Medicine & Medical Sciences 1989; 18:151-4.
57. Owen AW, Anderson TJ, Forrest AP. "Closed" biopsy for breast cancer. Journal of the Royal College of Surgeons of Edinburgh 1980; 25:237-41.
58. Place R, Velanovich V, Carter P. Fine needle aspiration in the clinical management of mammary masses. Surgery, Gynecology & Obstetrics 1993; 177:7-11.
59. Rotstein S, Nilsson B, Svane G, Gustavson-Kadaka E. Clinical examination, mammographic findings and cytological diagnosis in patients with breast disorders. Results of 9 years' follow-up. Acta Oncologica 1992; 31:393-7.
60. Saarela AO, Kiviniemi HO, Rissanen TJ, Paloneva TK. Nonpalpable breast lesions: pathologic correlation of ultrasonographically guided fine-needle aspiration biopsy. Journal of Ultrasound in Medicine 1996; 15:549-53; quiz 555-6.
61. Saunders G, Lakra Y, Libcke J. Comparison of needle aspiration cytologic diagnosis with excisional biopsy tissue diagnosis of palpable tumors of the breast in a community hospital. Surgery, Gynecology & Obstetrics 1991; 172:437-40.
62. Scopa CD, Koukouras D, Spiliotis J, et al. Comparison of fine needle aspiration and Tru-Cut biopsy of palpable mammary lesions. Cancer Detection & Prevention 1996; 20:620-4.
63. Sharifah NA, Hayati AR, Meah F, al-Rashid Z. Fine needle aspiration cytology of the breast: an analysis of 1,094 aspirates. Malaysian Journal of Pathology 1993; 15:53-8.
64. Sheridan WG, Davies JE, Lewis MH. Early experience with fine-needle aspiration cytology in a district general hospital. British Journal of Hospital Medicine 1987; 37:156-8.

65. Sneige N, Holder PD, Katz RL, et al. Fine-needle aspiration cytology of the male breast in a cancer center. Diagnostic Cytopathology 1993; 9:691-7.
66. Somers RG, Young GP, Kaplan MJ, Bernhard VM, Rosenberg M, Somers D. Fine-needle aspiration biopsy in the management of solid breast tumors. Archives of Surgery 1985; 120:673-7.
67. Stavric GD, Tevcev DT, Kaftandjiev DR, Novak JJ. Aspiration biopsy cytologic method in diagnosis of breast lesions: a critical review of 250 cases. Acta Cytologica 1973; 17:188-90.
68. Steinberg JL, Trudeau ME, Ryder DE, et al. Combined fine-needle aspiration, physical examination and mammography in the diagnosis of palpable breast masses: their relation to outcome for women with primary breast cancer. Canadian Journal of Surgery 1996; 39:302-11.
69. Strawbridge HT, Bassett AA, Foldes I. Role of cytology in management of lesions of the breast. Surgery, Gynecology & Obstetrics 1981; 152:1-7.
70. Svane G. Stereotaxic fine needle biopsy of nonpalpable breast lesions performed by the Mammotest. Recent Results in Cancer Research 1987; 105:95-6.
71. Teixidor HS, Wojtasek DA, Reiches EM, Santos-Buch CA, Minick CR. Fine-needle aspiration of breast biopsy specimens: correlation of histologic and cytologic findings. Radiology 1992; 184:55-8.
72. Thomas PA, Vazquez MF, Waisman J. Comparison of fine-needle aspiration and frozen section of palpable mammary lesions. Modern Pathology 1990; 3:570-4.
73. Ulanow RM, Galblum L, Canter JW. Fine needle aspiration in the diagnosis and management of solid breast lesions. American Journal of Surgery 1984; 148:653-7.
74. Verma K, Kapila K. The role of fine needle aspiration cytology of breast lumps in the management of patients. Indian Journal of Medical Research 1989; 90:135-9.
75. Vetrani A, Fulciniti F, Di Benedetto G, et al. Fine-needle aspiration biopsies of breast masses. An additional experience with 1153 cases (1985 to 1988) and a meta-analysis. Cancer 1992; 69:736-40.
76. Wallgren A, Zajicek J. Cytologic presentation of mammary carcinoma on aspiration biopsy smears. Acta Cytologica 1976; 20:469-78.
77. Wilson SL, Ehrmann RL. The cytologic diagnosis of breast aspirations. Acta Cytologica 1978; 22:470-5.
78. Wollenberg NJ, Caya JG, Clowry LJ. Fine needle aspiration cytology of the breast. A review of 321 cases with statistical evaluation. Acta Cytologica 1985; 29:425-9.
79. Young GP, Somers RG, Young I, Kaplan M, Cowan DF. Experience with a modified fine-needle aspiration biopsy technique in 533 breast cases. Diagnostic Cytopathology 1986; 2:91-8.
80. Yu SC, Lin PW, Liaw KY, et al. Fine needle aspiration cytology of breast tumors. Taiwan I Hsueh Hui Tsa Chih - Journal of the Formosan Medical Ass 1985; 84:727-34.
81. Zajdela A, Zillhardt P, Voillemot N. Cytological diagnosis by fine needle sampling without aspiration. Cancer 1987; 59:1201-5.
82. Zemba-Palko V, Minkowitz G, Kline TS. Occult breast lesions and aspiration biopsy: the Lankenau experience. Diagnostic Cytopathology 1994; 11:15-9.

Bibliography

New England Medical Center Personnel receiving pay from the research effort:

John B Wong
Nananda F Col
James E Stahl (now at the University of Pittsburgh)
Carol Nichols

TABLE OF CONTENTS

Diagnostic Strategies for Breast Cancer

I. Introduction

I.A. Overview of the Problem

Numerous studies have provided evidence that breast cancer screening by physical examination, mammography, or both can reduce mortality. In response, various national standard-setting bodies and professional societies have promulgated guidelines on breast cancer screening. Though there is general consensus that mammography screening over the age of 50 is indicated, screening between the ages of 40 and 50 has stirred controversy. In 1991, the National Cancer Institute conducted a Physician Visit Survey and found that 61% of all women over 40 had at least one mammogram over the prior two years. In this study, age-appropriate guidelines were defined as having a mammogram within the past two years if the woman were between the ages of 40 and 49, and within the past year if the woman were age 50 or older. Of note, compliance with screening guidelines declined with age, from a mammography screening rate of 68% among those aged 40-49 to 49% among women aged 50-64 and 40% of those 65 and older.[1] As one recent study though notes, five times as many cancers per 1000 first-screening mammograms were detected among women aged 50 or older than women under that age.[2]

Yet with women increasingly aware of the signs of breast cancer, diagnostic evaluations may become increasingly patient-initiated. A 1986 nationwide Access to Care survey discovered that women aged 20 to 39 years had the highest rate of clinical breast exams although no guideline recommendations support this practice.[3] This telephone survey also found that younger women had greater concerns about breast cancer, considered their personal risk as greater, and recognized the value of mammography in the early detection of breast cancer moreso than older women. The perception of personal risk seems greatest in the age cohort at lowest epidemiologic risk of breast cancer. Moreover, one community survey suggests that physicians are targeting the wrong women with mammography screening. In this 1991 study of two North Carolina counties, one quarter of women aged 30 to 39 years had a previous mammogram, and nineteen percent of physicians reported screening all women in this age range.[4] Both of these studies suggest a spillover effect from the impetus of increased breast cancer screening efforts.

These breast cancer screening efforts--the intended consequence of guidelines and the unintended spillover from them--lead to a cascade of diagnostic tests. Yet diagnostic tests are imperfect. They may fail to identify those with disease and thus give false reassurance (false negatives). Or tests may mislabel those free of disease as having the condition and cause undue anxiety (false positives). Each test raises different anxieties, exacts different costs, and imposes different risks of morbidity. Excisional biopsy and fine-needle aspiration cytology are both invasive tests, while mammography and the physical examination are not.

Current diagnostic modalities for the evaluation of breast abnormalities include principally: 1) breast clinical examination; 2) mammography; 3) fine-needle aspiration cytology (FNAC); 4) various types of biopsy procedures--core, Tru Cut, and excisional. Several technologies, such as ultrasound and mammography, assist in localizing lesions for biopsy. Apart from localization, ultrasound may play an adjunct role in sorting cystic from solid masses for FNAC.

These tests for evaluating breast cancer must be used in combination and in sequence to reach a diagnostic endpoint. At each step of this diagnostic pathway, the patient presents with a certain

pre-test probability of disease, shaped by antecedent history and findings. Each test and testing sequence yields a characteristic sensitivity and specificity. Our overview of the literature suggests that: 1) each diagnostic test presents a trade-off between sensitivity (true-positive rate) and specificity (1-false-positive rate), between true positives and false positives; and 2) where this trade-off occurs depends often on modifiable factors, such as training in breast physical examination, standardizing the reporting of mammography, or setting minimum standards for competency in FNAC or excisional biopsy performance. So if we can identify the optimal operating point for a diagnostic test, where the trade-off in sensitivity and specificity is best, then we can develop interventions to calibrate the performance of these diagnostic tests to reach that optimum operating point. This two-year research project proposes to identify the optimal diagnostic test strategies.

I.B. Purpose of Present Work

To identify the optimal diagnostic test strategies for evaluating breast abnormalities, we proposed a multi-step approach:

- To perform a comprehensive and structured literature review of diagnostic tests for the evaluation of breast cancer and to apply quantitative meta-analysis, if appropriate, in order to derive estimates of test characteristics, complication rates, and outcomes.
- To construct a decision analysis that evaluates the optimal diagnostic testing strategy for breast cancer evaluation, assesses the magnitude of misclassification, and highlights the limitations in currently available data.
- To compare the incremental cost-effectiveness and misclassification costs for each diagnostic testing strategy for particular clinical presentations leading to breast cancer evaluation.
- To conduct focus groups of primary care physicians, referral physicians, and patients in order to give this work expert review and feedback.

Within this framework, we have focused on the structured literature review and meta-analysis in project year 1. Existing literature reviews have filled in summary estimates for current diagnostic modalities, except for fine-needle aspiration cytology. We decided to devote our attention to this pivotal procedure in the diagnostic work-up of breast abnormalities. We have several reasons for taking this strategy:

1. FNAC sits at the center of the diagnostic testing sequence in evaluating breast abnormalities. As a procedure, it is considered less definitive than excisional biopsy. Thus some have questioned whether it is cost-effective to use it in the diagnostic sequence.
2. Others would argue that the FNAC, in combination with clinical breast examination and mammography, obviates the need for pursuing excisional biopsy, which usually requires referral to a surgeon.
3. FNAC also raises the challenges of clinical privileging. Its test characteristics are likely to be dependent on operator performance. Whether the quality of performance is volume-related, exhibits a practice effect and plateaus, or requires a specialty clinic is debatable. A decision analysis might describe the challenge region (that is, the minimum threshold of test performance) to which an operator must perform to be "competent" in the use of FNAC.

As this first year comes to a close, we are also conducting physician focus groups to explore the diagnostic decision making process in the evaluation of breast abnormalities. In addition, we have

sought and received IRB approval to field a patient survey instrument to elicit preferences and perceptions of women undergoing diagnostic evaluation. Both the physician and patient feedback started with a series of expert and lay interviews which laid the groundwork. Finally, we have also begun the process of outlining the diagnostic pathways that will become the branches of a decision tree analysis.

II. Narrative

II.A. Methodology

A.1. Structured Literature Review

Through computerized literature searches of MEDLINE, we have recruited relevant journal articles on fine-needle aspiration cytology. To accomplish the computerized search, we used the Grateful Med interface with MEDLINE and imported references into a bibliographic retrieval program, Endnote Plus/Endlink. Within this software program, we were able to track and record the status of articles in the review process, as well as to sort the database of articles by keywords into sub-libraries. This program enabled us to maintain an up-to-date registry of the articles accepted or rejected, the reason for the decision, and other pertinent information in the review process, as well as bibliographic information. We retrieved journal articles from local biomedical libraries and through inter-library photocopying requests.

Search strategy. For the computerized literature search, we crossed the MESH terms BIOPSY, NEEDLE and (BREAST or BREAST DISEASES) on MEDLINE. As we narrow the number of articles to those finally accepted, we plan to identify fugitive literature by 1) ancestral tracing of bibliographic references and 2) following up citations suggested by experts. Telephone requests to professional medical societies (ACOG and ACR) have not yielded alternative reference listings for these articles.

Dates included. We searched the MEDLINE database between the years 1966 (the year electronic cataloguing began) and 1994. We considered criteria that might set a later search date (e.g., a technologic advance that would render older literature findings obsolete). However, no such criterion suggested a compelling change in technology, and so we decided to opt for a broader search.

Delimiters. We applied a multi-step process to limit our search. First, we accepted only articles written in English. Given practical considerations, translation would not be feasible. However, we did not exclude studies conducted in foreign countries. Secondly, we eliminated articles identified as non-original contributions in the MEDLINE search (e.g., letter, comment, case report, review, news, or editorial). Though it is possible that an occasional review, letter or editorial might introduce new data, the format would not typically allow the data abstraction required for inclusion in a meta-analysis.

Exclusion criteria. We established a set of exclusion criteria applied first to the literature abstracts and later to selected articles pulled for further review. These criteria include:

NR	Not relevant	Despite the match with MESH terms, some articles do not present data on the test characteristics of breast cancer work-up. Others mention breast cancer evaluation only

		incidentally, and therefore, are also not relevant to this meta-analysis.
IS	Insufficient sample	In some studies, a subject underwent more than one test. In determining sample size, we took the number of tests, as opposed to the number of subjects, as the unit of analysis. Though somewhat arbitrary, the cutpoint we set was N<100. We can offer a back-of-the-envelope justification for this cutpoint.* Moreover, we flagged studies with insufficient sample, so that we could return to assess the effect of studies with small numbers on our meta-analysis results.
NO	No original work	Excluding by MESH terms articles without original data (e.g., editorial, reviews, correspondence), the capture was still incomplete, and at the abstract and full-text journal review level, we flagged other articles as having no original data.
SP	Special population	By study design, the external validity of some articles was limited to a subset of the population. For example, the subjects might all have familial predisposition to breast cancer or already have had breast cancer once.
AR	Absent reference or gold standard	Test characteristics are only meaningful when measured against a known reference standard. In most cases of fine needle aspiration cytology, we take biopsy and clinical follow-up as the reference or gold standard. For clinical follow-up, we decided against requiring a follow-up period of specified duration as qualifying.
PV	Procedural variation	Some articles focus on an innovative or experimental technique, such as the immunocytochemistry or receptor status of tissue samples. They do not replace the diagnostic test (e.g., mammography, FNAC, biopsy, breast examination), but may be used in conjunction with one.
SS	Special subset	Some studies focus only on a specific type of breast tissue or neoplasm. The denominator under study is restricted to accessory breast tissue, lobular carcinoma or a specific tissue type. Test characteristics derived from such articles do not apply necessarily to all comers presenting in a patient population.
VB	Verification bias	By verification bias, we refer to the inconsistent application of the reference test or gold standard. This leads to biased

* $N = \frac{z^2 p(1-p)}{L^2} = \frac{1.96^2 (0.9)(0.1)}{(0.05)^2} = 138$, where N=sample size, L=half width of confidence interval, p=sensitivity or specificity sought

		estimates of sensitivity and specificity.
OT	Other	Other unanticipated reasons might prompt exclusion of articles, and these articles warrant further examination by the investigators on a case-by-case basis.

We used a disjunctive positivity criterion for exclusion, that is, an article could be rejected by the first exclusion criterion flagged. With some articles excluded on the basis of the abstract and others on a complete reading, we did not find it practical to identify all exclusion criteria by which an article might be rejected. Consequently, we decided that the exclusion criteria selected for rejecting an article does not have to be the same for the two independent reviewers, as long as both reviewers agree that an article should be rejected.

Two reviewers read each of the abstracts and identified by code the reasons for exclusion, if any. The ratings by each reviewer were calibrated through a training set of 70 abstracts evaluated by the Principal Investigator. Where differences arise between two reviewers, these were discussed with the Principal Investigator, and if any question remained, the full-text journal article was retrieved for further review. We assessed inter-rater reliability between reviewers in their decision to accept or reject articles on the basis of their abstracts.

Data abstraction form. After many revisions, we piloted and fielded a data abstraction form suitable for the purposes of our meta-analysis (see **Appendix I**). The print version has incorporated key data elements such as: 1) identifying information of the journal article; 2) the diagnostic test(s) under study, equipment used, and localization techniques applied; 3) the sample size of tests performed and subjects recruited, with note made of exclusions and dropouts; 4) type of population under study (inclusion and exclusion criteria for subjects); 5) diagnostic test characteristics reported; 6) provider or operator experience; 7) facility where tests were performed; 8) complications; and 9) study design, such as cohort or case series, and data collection approach. We also collected information that bears on the quality of the evidence, such as the reference standard and clinical follow-up used. This allowed us to ascertain verification bias in the studies. We have generated a form designed to accommodate the range of diagnostic tests and testing sequences in the literature.

Data abstraction process. Two reviewers perform data abstraction from each full-text journal article. The Principal Investigator and a co-investigator trained all reviewers, and we discussed their data abstractions on a training set of articles. At this review stage, we perform a second screen applying the established exclusion criteria, and reviewers record whether they accept or reject the article under consideration. Using raw data reported in the study, the reviewer also recalculates the test characteristics. This is necessary since we have noted that studies vary in their interpretation of atypical or inadequate test results. Sometimes they figure into the calculations of the original study's test characteristics, and sometimes they do not. Reviewers fill out these data abstraction forms as completely as possible and bring points of contention or confusion to the Principal Investigator and co-investigators. Of course, the forms also flag missing data.

Construction of evidence tables. Data abstraction forms from both reviewers are then entered into a customized computer database that we constructed in Microsoft Access. The database permits the flexible generation of evidence tables that compare findings across studies included in the meta-analysis. These side-by-side comparison charts enable bivariate analysis, such as the influence of study sample size on sensitivity or the relationship of age to specificity. For graphical display work

and some calculations, we export from the Microsoft Access database to a Microsoft Excel spreadsheet program.

Quality control measures. Data entry permits a cross-check on the reliability of the data abstraction process. Where discordance is noted, it can be resolved by 1) referring to the original full-text article; 2) discussion between reviewers; and 3) when necessary, adjudication by the Principal Investigator and his co-investigators. In addition, the Principal Investigator and co-investigators are conducting a quality check on a 10% randomly selected sample of the full-text journal articles pulled for review.

A.2. Meta-analysis

Meta-analysis is a quantitative approach to combine data from multiple studies on the same topic. In this phase of our project, we are both generating summary estimates of the diagnostic test characteristics and examining how these test characteristics are influenced by heterogeneity in study design.

Test characteristics. Test sensitivity is defined as a probability, the p[positive test result | disease], and test specificity, as the p[negative test result | no disease]. These test characteristics apply to binary outcomes, but fine-needle aspiration cytology does not always yield binary results.

In fact, most studies report atypical and inadequate findings as well as positive and negative. Each category deserves separate consideration. Atypical findings on FNAC raise the level of clinical suspicion and often are treated as positive in the clinical setting, insofar as clinicians are inclined to investigate these findings further. However, in some settings, atypical findings register a level of uncertainty on the part of the cytopathologist. By including atypicals in the calculation of sensitivity or specificity, we might be distorting the calculations of test characteristics. We have opted to calculate test characteristics both with and without the inclusion of atypical cases as positive.

Most studies also report inadequate FNAC samples in varying proportions. Inadequate samples result from factors such as operator experience, number of aspiration passes, localization mode, and size of the tumor. In contrast to atypical samples, inadequate samples represent a problem of feasibility as opposed to diagnostic accuracy. Some series have cytotechnologists reviewing FNAC samples at the time of the procedure in order to determine if they are acellular and warrant immediate repeat aspiration. Typically studies exclude them from the analyses, and we too have excluded them from the calculation of test characteristics. However, we also examine graphically the relationship between the percentage of inadequate samples and the reported test characteristics.

Publication bias. Publication bias occurs when studies appear in the literature only if they are well-conducted or offer statistically significant results. Consequently, performing meta-analysis only on these published studies might yield biased estimates. Light and Pillemer describe a quasi-statistical, graphical technique for assessing publication bias. The funnel plot[5] graphs the effect measure on the x-axis, and the sample size, on the y-axis. Absent publication bias, the funnel plot should take the shape of a funnel with the large opening down and the narrow end up, centered over the true effect size. The funnel shape results from the expected sampling variability across studies. For this meta-analysis of a diagnostic test, we plot the test characteristic rather than the effect measure on the x-axis.

Verification bias. We took as the reference standard both excisional biopsy and clinical follow-up. The failure to pursue a diagnostic work-up to its gold or reference standard would otherwise miss false negatives. If a study reported more than 20% loss to clinical follow-up, then it would be excluded from our meta-analysis.

The duration and nature of clinical follow-up varies across studies. The literature estimates tumor doubling time at 100 days with a range from 30-200 days.[6] How does this variability in clinical follow-up time influence test characteristics, particularly the false-negative rate? To assess this issue, we would have calculated the mean duration of clinical follow-up and have plotted clinical follow-up against FNAC test characteristics. However, absent information about mean duration, we could consider the range of clinical follow-up and plot the low end of the range against test characteristics.

Selection bias. Through our exclusion criteria, we have removed studies that focus on a particular patient subgroup or breast tissue type. Studies focusing on specific patient subgroups may have selected for patients with a family history of breast cancer or a recurrence of breast cancer. In studies focusing on a specific breast tissue type, inclusion is based retrospectively on the histologically confirmed outcome of the FNAC, such as cysts or colloid carcinoma of the breast.

However, other factors shape whether the patients at study entry are at relatively higher or lower risk of having breast cancer diagnosed. What changes is the pre-test probability, and in turn, the post-test probability of disease after the diagnostic test is used. These factors include: 1) positive findings on tests preceding study enrollment; 2) palpability of breast lesion; and if reported, 3) size of breast lesion discovered. To evaluate the influence of these factors, we have compared the derived test characteristics of studies that differ along these dimensions. For example, we examine whether studies looking at palpable breast lesions report a higher sensitivity or specificity than studies looking at nonpalpable lesions.

Some studies apply a battery of tests. When used in parallel, a battery of tests does not take advantage of the changes in pre- and post-test probability that accrue from learning of each diagnostic test result serially. Done serially, the diagnostic tests may no longer be considered independent. Increasingly, clinics try to provide rapid diagnostic work-ups for patients with identified breast lesions. By doing so, they may resort to ordering tests in parallel rather than serially. For example, the mammogram and the FNAC are done in combination rather than in succession. In our meta-analysis, we plan to study whether test characteristics systematically differ when done in parallel or in succession.

Patient population. The populations recruited to these studies differ along factors like age. By recording the demographic characteristics of the patients under study, we can analyze whether age has a significant influence on FNAC results.

Studies may differ considerably in the ratio of benign: malignant lesions (B:M) discovered.[7,8] Alternatively, this B:M ratio may be represented as the percentage malignant. Some have suggested it might be used to counsel women about their breast cancer risk at the time of biopsy, but others have argued that it might be used to judge the adequacy of care given by a physician or hospital. However, variability in the B:M ratio may point to failings in its use as a quality of care indicator. Several factors may contribute to this variability. One source is differences in the patient population presenting for evaluation, and these factors include age, race, socioeconomic status, temporal trends in the awareness of breast cancer. Another source traces to non-population-related

factors such as the greater use of diagnostic techniques, improvements in the prebiopsy screening of breast lesions, or differences in the definition of benign breast lesions or of cancer. Though the source of variability casts a wide net, we compared studies reporting a high prevalence of malignant lesions (as represented in the benign:malignant ratio) in their test results against the rest of the studies.

Testing site. We also recorded the site where FNACs were performed and compared test characteristics, the number of atypical samples, and the number of inadequate samples obtained across these sites. These practice settings could potentially range from the generalist's office to the tertiary referral center specializing in breast evaluation. Of course, the methods section of these papers may lack the detail necessary to classify the practice setting accurately. If misclassification results, the bias towards a negative finding is increased. Underlying differences in practice settings, of course, may be the experience of the operator and the referral pattern of patients.

Secular trends and FNAC technique. As a procedure, the technique of FNAC has remained quite constant over time. However, there may have been changes in equipment (e.g., needle size), in the presenting patient population, and in the experience of operators. Insofar as studies report the FNAC needle size in their methods, we compare differences among these studies. We also plot temporal trends in the reported sensitivity and specificity of FNAC. Though the year of recruitment would have been ideal for this purpose, we have taken the date of publication as proxy for this initial analysis.

Operator technique. Some studies report the number of passes made by the FNAC operator. Presumably the greater the number of passes, the greater the yield for diagnosis. How does this influence sensitivity or specificity? Through our meta-analysis, we analyze this relationship between the number of passes and the reported test characteristics.

Operator experience. For various procedures, the health services research literature suggests a relationship between volume and quality.[9] Operator training and experience in FNAC has been reported to result in a higher level of test sensitivity and specificity.[10] Similarly, for fine-needle aspiration biopsy, we assess the influence of operator experience on this test's characteristics. However, such experience is seldom directly reported in the literature. Other variables though may serve as proxy, and we have considered several indirect approaches.

- *Early vs. established practice.* Studies sometimes characterize their procedural experiences as early, even their first cases, or established. Such data allows for comparisons between novices and the experienced in different studies.
- *Early vs. later practice experience.* A few articles may tackle this issue by contrasting the earlier and the later experiences of a center performing fine-needle aspiration cytology.
- *Number of operators.* The greater the number of operators, the lower the distributed volume of procedures among them. In addition, when there is more than one operator, variability in individual performance may contribute to an apparent relationship between volume and quality.
- *Volume per unit time.* By using the accrual period as the denominator, the number of procedures during that period might provide a proxy for volume. If the number of operators is noted, then we can also calculate the number of procedures per unit time. Moreover, the volume per unit time could also take into account the number of operators as well and yield the number of procedures per operator per unit time.

Method for combining estimates of test characteristics. Statistical approaches used for the meta-analysis of randomized controlled trials, such as Dersimonian and Laird or the Peto method, do not apply to the meta-analysis of diagnostic tests. Several methods have been advanced for combining estimates of diagnostic tests. One approach is called "collapsing." The test characteristic is expressed as a proportion, and the combined estimate takes an average of these study proportions, each weighted by the respective study sample size. The disadvantage to this method is that it ignores among-study heterogeneity in the calculation of variance.

A modification of the collapsing approach comes from the survey sampling literature. As Berlin, et al. have described, "in this literature, a study is viewed as a naturally occurring 'cluster' of individuals. The point estimate of the combined sensitivity or specificity under cluster sampling is the same as that used in collapsing . . ."[11] Cluster sampling provides an unbiased estimate and accounts for among-study heterogeneity. Using a Fortran-based program written for this purpose, we derive estimates of test characteristics by using the cluster sampling methodology.

Statistical test for study heterogeneity. Typically, the results of studies in any clinical area are not completely uniform. Random (sampling) error would generally lead to a certain amount of variability, or heterogeneity, among study results. We used a formal test of this heterogeneity to address the question of whether the observed variability among study results is consistent with what one would expect by chance (i.e., due to sampling error alone) or if it exceeded random variation. The null hypothesis for the test is "homogeneity," or equality of estimates across studies. Rejecting the null implies that at least one of the studies is estimating a different parameter than the others. This might result from the sensitivity, e.g., being different in a study that examines a different subgroup of the population of interest. Failure to reject the null does not necessarily imply equality of estimates, however, because of the well-known poor statistical power of the test. Formally, the test, called the Q-statistic, is a chi-square with degrees of freedom equal to one less than the number of studies being combined.

When significant heterogeneity is found, one can calculate summaries for different subgroups of studies. However, this analysis must await completion of the structured literature review. In addition, heterogeneity may result from the use of different cutpoints in various studies to determine whether a test is positive or not. Should this be the case, we can resort to transforming the sensitivities and specificities from each study into a common receiver operating characteristic curve.

A.3. Decision Analysis

Groundwork for decision analysis. Our work on the structured literature review and meta-analysis has set the stage for the follow-on decision analysis. The structured literature review has helped us to identify the key diagnostic pathways and contemporary controversies related to them. Through the review, decision alternatives are outlined along with their clinical outcomes. The meta-analysis has focused on a diagnostic test central to this pathway--fine-needle aspiration cytology. By recognizing the limitations of individual studies and of their meta-analysis, we gain a sense of where a decision analysis may begin to help pull together what we do know. We are developing an evolving map of these decision pathways as a first step. However, an important aspect of this work has been eliciting feedback from both physicians and patients.

Physician feedback At this stage, we have planned the hosting of two physician focus groups. They provide reality testing on clinical decision making involving breast abnormalities. Highlighting some of the issues entering the planning process, we deliberated over the following:

- *Composition and recruitment.* The physician focus group brought together an interdisciplinary panel of practitioners and referral specialists. They would be drawn from different backgrounds to share their perspectives: general internal medicine/family practice, gynecology, surgery, oncology, radiology, and cytopathology. One group consisted of physicians largely from one university referral network, while the other draws participants more broadly from the larger community.
- *Presentation.* To deal with the multitude of clinical factors, we designed a matrix that decomposed the clinical vignette into the component decision factors. We presented these issues with the assistance of an overhead projector and an audience response system using handheld keypads. The audience response system enables us not only to record the answers given in each focus group that we conducted, but also to compare across groups. A pre-focus group survey gauged the referral patterns, practice volume, diagnostic pathways, and other aspects of practice among the various focus group participants.
- *Defining the issues.* The topics for focus group discussion were developed with input from the research team, a series of individual expert interviews, and our literature review. We sought an understanding of 1) the clinical factors for using FNAC, 2) the sequencing of diagnostic tests, 3) local variations in work-up, evaluation, and test interpretation, 4) the role of patient expectations and concerns, and 5) the emergence of new diagnostic techniques, such as core biopsy. More specifically, the physician focus group offered an opportunity to learn how clinicians consider various conundrums such as: 1) when to repeat FNAC or opt for biopsy; 2) what type of presentation might obviate the need for more extensive diagnostic work-up; 3) how the suspicion that a breast mass is cystic influences the subsequent diagnostic pathway; or 4) the interpretation of atypical and inadequate samples on FNAC.

Patient feedback We have also built the discussion framework for eliciting preferences and attitudes from patients. This input will eventually contribute to the shaping of utilities in the decision analysis. For a variety of reasons, this process remains more open-ended in design at this point. These reasons reflect the following issues:

- *Diverse patient populations.* We hypothesize that the perspectives of patients change as they proceed through the diagnostic work-up. Patients with diagnosed breast cancer would likely weigh utilities from diagnostic evaluation differently than patients who experience false-positive work-ups. Recruiting a patient focus group presents greater difficulties. The composition of the focus group does not easily divide into categories, like it did with physician specialty types. Natural recruitment sources, such as breast cancer support groups, carry a particular bent to these issues. Particular clinical settings, such as the waiting room for mammography, select for patients at a similar point in the diagnostic work-up. And the participation of breast cancer patients in these focus groups may introduce undue influence on the group discussions.
- *Instrument design.*To accommodate this range of patients, we developed an instrument that focused on key issues identified in interviews with social workers, patients, and the research team. In addition, we drew upon themes noted in the literature on the psychological and behavioral effects of diagnostic evaluation for breast cancer. To obtain utility estimates, we will have to pilot and conduct individual patient interviews using the time-tradeoff or similar

technique. Similarly, the instrument can provide a framework for patient focus group discussions if we find that approach productive.

We have submitted a successful IRB request for conducting patient interviews in various clinical settings. By doing so, we have addressed and provided reassurances about the confidentiality of patient data, the informed consent process, and related issues.

II.B. Results and Discussion

B.1. Structured Literature Review

Using MeSH subject headings and searching the years 1966 through 1994, we identified 1959 potential journal articles on fine-needle aspiration cytology and related diagnostic procedures, such as breast biopsy. By limiting the search only to articles written in English, we reduced the number of candidate articles from 1959 to 1575. We were also able to exclude articles categorized as a non-original contribution (n=410) and articles focusing on needle biopsy in another organ system (n=43). This left a total of 1122 articles for abstract review.

Type	Total number of articles deleted
Letter	134
Comment	59
Case Report	113
Review	81
News	5
Editorial	18
	410
Thyroid	8
Lymph node	8
Bone marrow	10
Liver	8
Lung	4
Prostate	1
Salivary gland	2
Neck	1
Abdomen	1
	43

The Principal Investigator and research assistant reviewed 1122 abstracts. Of these, 467 journal articles (42%) were accepted, and 655 (58%) rejected using our exclusion criteria. The percentage agreement between both reviewers reached 94% (430 accepted with concordance of both reviewers, 37 without; 624 rejected with concordance, 31 without).

By abstract, we then sorted accepted studies into categories of those dealing primarily with FNAC (363), core biopsy (15), Tru-Cut biopsy (7), excisional biopsy (87), mammography (17), and ultrasound (2). Despite using more than one biomedical library in the region, we still have 74 interlibrary photocopying requests outstanding. However, 80% of the FNAC studies have already been retrieved and reviewed.

By dropping articles when the first exclusion criterion is flagged, we cannot ascertain the most common reasons for exclusion. However, the abstract reviewers paid particular attention to making note of articles excluded on the basis of sample size. In our database, we excluded only 11 articles because their sample size did not exceed 100 diagnostic tests.

Two independent readers reviewed and abstracted data from each accepted, full-text journal article. Of 363 accepted FNAC articles, 289 have been retrieved, and 252 (69%) reviewed by at least one reader.

Number of Articles Accepted for Meta-Analysis: 100 (40%)
Number of Articles Rejected for Meta-Analysis: 152 (60%)

Thus far, 168 articles have been fully reviewed by two readers. We track this process through our bibliographic database on Endnote Plus. From this database, we have generated as an example a selected listing of the first 100 articles accepted to the meta-analysis (see **Appendix II**). Once all articles have been reviewed by two readers, we can use concordance measures to examine their decisions to accept or reject full-text journal articles. However, the primary purpose of the two reader system is quality control for the multiple data elements in the abstraction process, e.g., the sample size and test characteristics. A second layer of quality control measures will involve review of a 10% randomly selected sample of the full-text journal articles. We have used a random number generator program to identify these articles and have tracked this sample through the reference number assigned by our Endnote Plus bibliographic retrieval program.

We can assess reliability of the two-reader abstraction process during database entry into Microsoft Access. This computerized database enables us to produce evidence tables, and this provides us a way to view study findings side by side. The charts noted throughout the *Results and Discussion* section (see **Appendix IV**) draw upon data exported from Microsoft Access into Microsoft Excel for graphical display.

II.B.2. Meta-Analysis

At this juncture in the project, we can provide an overview of results to date. By doing so, we delineate the framework for our analytic strategy. However, the findings on fine-needle aspiration cytology are incomplete and preliminary. We present the results as graphical displays, a steppingstone to subsequent work using regression analysis and other statistical methods. We draw upon the Microsoft Access database which currently has 101 of the abstracted studies in its repository.

Test Characteristics

Many studies reported atypical results for FNAC, and these were variably counted as positive or negative tests in calculating sensitivity and specificity. Of 101 studies, thirty-six reported atypicals as a separate category, while fifty-four studies did not report any atypical results. Nine had included them in the test positive category of suspicious for malignancy, and two studies included them in the test negative category of benign findings. We performed our calculations of test characteristics both with and without the atypical results included in the test positive category.

Viewing the scatter plot of the percentage of reported atypicals versus the test characteristics, a slight downward trend in sensitivity (Chart 1) may be associated with an increasing percentage of atypical results. No such relationship appears in the graphical plot for specificity (Chart 2). In

these charts, the test characteristics are calculated with atypical results counting as positive test results. Subsequent analyses may depict the relationship between atypical findings and FNAC test characteristics. The thresholds for calling a sample atypical as opposed to positive or negative have direct impact on the diagnostic yield of FNAC.

Most of our studies reported inadequate samples, and their influence on sensitivity and specificity is not direct since they are not included in the calculation of the test characteristics. But one might propose that they exert an indirect influence over test characteristics in the sense that they are operator and technique dependent. Therefore, a study with many inadequate samples might have poorer test characteristics in general. Alternatively, a study with a high percentage of inadequate samples might reflect a higher quality threshold demanded for cytopathology reading. In the graphs plotting inadequate samples against sensitivity or specificity, no clearcut relationships are seen yet (Charts 3 and 4).

Publication bias

In order to assess publication bias, we have graphed two funnel plots (Charts 5 and 6), with sample size on the y-axis and test characteristics as the "effect" measure on the x-axis. There were sixteen studies which had sample sizes over one thousand, but as these were outliers, they are not depicted on these charts. By amplifying the area displaying studies with sample sizes 1000 and under, one can better appreciate whether a funnel shape pattern to the plots is seen or not. There is somewhat wider dispersion of study test characteristics at the bottom of the funnel, where the sample sizes are smaller. This can be attributable to the anticipated greater statistical variability that comes with smaller sample sizes. If there had been a pronounced publication bias, we would have expected that even studies with smaller sample sizes would only have published results with near perfect sensitivity or specificity. However, this does not appear to be the case here.

Verification Bias

In our meta-analysis, both biopsy and clinical follow-up were accepted as reference standards. A total of 30 articles contained biopsy and clinical follow-up data, while 71 articles used only biopsy as the gold standard. Only a subset of the studies reported the duration of clinical follow-up, and most of them did not provide a mean or median. So we focused on the ranges given for clinical follow-up time. They varied from a minimum of 3 months to over 36 months, but the majority (27/30) of the studies were in the range of 6-12 months (Chart 7).

Presumably a longer follow-up period might permit the detection of latent or missed malignancies, or in other words, the false negatives. Of course, there also comes a point when the follow-up period is so long that malignancies cropping up were not missed at the point of initial diagnostic evaluation, but rather of new onset during the follow-up period. In looking for a relationship between follow-up duration and test characteristics, we did not find any discernible trend. This may be due to many factors and may trace to the quality of the follow-up, what additional diagnostic tests were used for follow-up, or patient compliance with follow-up. Many studies do not indicate how many patients, if any, were lost to follow up.

Absent data on the mean clinical follow-up duration, we plotted the lower end of the clinical follow-up range as a proxy against the false-negative rate (Chart 8). Though the chart suggests more false negatives detected with longer clinical follow-up, these findings are quite preliminary.

Selection Bias

The diagnostic tests performed on the patients before entering a study can alter the pre-FNAC probability of detecting a malignancy or a benign lesion. For example, many patients who were included in the studies had suspicious mammographic lesions or palpable masses on clinical exam (which implies the lesion had attained a certain size to be palpable). These prior test findings affect the pre-test probability or pre-test odds. Though sensitivity and specificity are theoretically invariant, these test characteristics can become biased by the resultant spectrum bias if only selected cases are referred on for further diagnostic work-up.

Of the studies included so far for meta-analysis, 25 studies noted a clinical breast exam as the pre-study test; 32 received a mammogram as well as the clinical exam, 25 had a mammogram only, and others had larger batteries of tests that included ultrasound and thermography.

Typically breast lesions are characterized as either palpable or nonpalpable. Of course, the nonpalpable lesions tend to be mammographically detected and smaller in size. Most studies did not report the size of detected breast lesions, so palpability will have to be taken as proxy. There are 22 studies of nonpalpable lesions, 44 of palpable lesions, and 7 of both. In Charts 14 and 15, we have generated a histogram comparing the pooled or summary test characteristic by pre-FNAC diagnostic work-up. However, to date, these differences in sensitivity and specificity are not statistically significant in pairwise comparisons between clinical exam or mammography and the combined diagnostic strategy of clinical exam plus mammography. Whether this implies conditional independence among tests done serially or in parallel requires further analysis.

Patient Population

Only 21 studies report the mean age of patients, and among these, they ranged from 38 to 65 years of age. A scatter plot of mean patient age versus test characteristics did not reveal any pattern; however, subsequent analyses will need to be performed. For example, we might compare those studies with a mean age greater than 50 with those reporting a mean age below 50. Of course, we risk a negative finding since characterizing a study population by its mean age rather than by the age of its individual subjects diminishes the power to detect an age-related difference in sensitivity or specificity of FNAC.

Studies also differ considerably in the benign : malignant ratio achieved, and this suggests a different operating threshold for pursuing FNAC. If only very suspicious lesions are being sent onto FNAC, then the benign:malignant ratio will decrease, and vice versa. In Charts 10 and 11 (test characteristic vs. benign:malignant ratio), we can see that the benign : malignant ratio remains mostly in the 1-4 range, but with still a rather wide dispersion of diagnostic yield. Further analyses will be necessary to discern whether there is a significant pattern to these plots.

Testing Site

Of the studies accepted for meta-analysis, 16 were regional or general hospitals, 15 were highly specialized centers like the Karolinska Institute of Stockholm or the Mayo Clinic, 40 were university centers, and 29 centers did not give enough information to judge their characteristics. If testing sites can be better characterized, we will examine this more closely.

Secular Trends and FNAC technique

The needle gauges used in FNAC vary from center to center or even from operator to operator. In our study, the gauges used ranged from 14 to 25, with the majority of the centers using 21-22 gauge needles. When plotting the effect of needle size on test characteristics, we see the greatest

variability is present at the most common gauge (22) and that sensitivity and specificity improving with either larger or smaller needle sizes (see Chart 19). This is probably due to two factors: 1) as needle size increases, it becomes progressively easier to get a good sample, and 2) the smaller needle gauges were used in very few studies with correspondingly less variability seen. Further analyses might evaluate the percentage of inadequate sampling against needle size.

To study temporal trends in the reported sensitivity and specificity of FNAC, we have used the date of publication as a proxy for year of recruitment (see Chart 20). As FNAC has become more accepted over time, more studies have been published. Interestingly, there is a wider spread in derived test characteristics with more recent studies. One explanation might be that earlier studies might exhibit greater publication bias or were conducted by investigators who pioneered or were more experienced in the technique. Also with the advent of screening mammography, the target lesions being aspirated may have progressively grown smaller. All these factors could explain this trend.

Localization technique also influences the test characteristics. With mammographically detected nonpalpable lesions, FNAC needs to be performed under some kind of localization technique. The most common localization mode used in our review was a stereotactic device based on X-rays or palpability in the case of palpable lesions. Charts 16 and 17 show preliminary pooled estimates of sensitivity and specificity comparing those studies using palpation against those studies using stereotactic localization. Within the bounds of confidence intervals, these two localization approaches have similar test characteristics; however, the underlying lesions may be different. In subsequent work, we will try to compare these two localization techniques in studies that focus solely on palpable lesions.

Operator technique

As the number of passes taken on FNAC increases, the variability in the reported test sensitivity decreases (see Chart 18). Whether there is a trend towards improved sensitivity is not clear from this initial graphical depiction.

Operator experience

Very few studies mentioned explicitly the number of aspirators, except in the cases where the aspirator was the author of the paper. The majority of the articles either did not mention the aspirator at all or implied the presence of several. Using the data from the articles that did contain this information, we plotted the test characteristics against the number of aspirators. The graphical display suggests that single aspirators had better sensitivity and specificity than the multiple aspirators. It is possible that single operators tended to be experienced, and the multiple operators may have included trainees. However, as we complete the data abstraction, we will return to calculate pooled estimates of sensitivity and specificity for studies with one, two, or more aspirators.

When the data were available, we calculated volume per operator per unit time by dividing the number of total aspirations by the number of aspirators by the study duration (in months). Only 25 studies had the required data. Charts 12 and 13 show these plots of test characteristic vs. volume per operator per unit time. On each chart, there is a wider dispersion of sensitivity or specificity among those studies where aspirators performed a lower volume of procedures. At this point though, we cannot tell whether this results from greater statistical variability with the smaller sample sizes per aspirator or points to a volume-quality relationship.

We do not report combined estimates of the test characteristics for all studies entered into the meta-analysis database to date. These calculations will be performed when the structured literature review and data abstraction processes are completed. Similarly, the statistical test for study heterogeneity will be applied then as well.

II.B.3. Decision Analysis

To lay the groundwork for the decision analysis, we have conducted a series of over 9 interviews with expert physicians involved in the field of breast cancer. These interviews have provided a clinical picture of the diagnosis of breast cancer and set the framework for several focus groups. Also the interviews shed light on unpublished, ongoing, or recently published studies which have not been detected by our MEDLINE search. The physicians interviewed comprised well respected members of different specialties involved in the diagnosis and treatment of breast cancer, e.g., a gynecologist involved in a novel Breast Cancer Risk Evaluation Program, a radiologist involved in investigation of MRI, a breast surgeon who uses FNA frequently but who is changing over to core biopsy, and a vice-chair of pathology who is involved in a study comparing FNAC to core biopsy. These interviews not only gave a current clinical picture but also indicated how rapidly changing these diagnostic strategies can be.

The focus groups were designed to investigate the clinical criteria used by the different specialists involved in breast cancer diagnosis in all its stages, and to challenge them and seek out areas of controversy and consensus. Participants in the groups comprise representatives from all specialties involved in the care of such patients: gynecologists, radiologists, internists, pathologists, oncologists, and surgeons. Clinical vignettes served to challenge the already established criteria for diagnosis of breast cancer. We present here only some of the findings from the first physician focus group.

The physician focus groups served to clarify diagnostic trends among different specialists. For example, surgeons tend to be more apt to excise all benign lesions, while the radiologists recommend excision if it is a palpable lesion or fibrocystic change only. Many clinicians, once they palpate a nodule, will send the patient to mammography irrespective of their age. However, the radiologists were adamant in the use of ultrasound in women under 35, and even in women over this age. On other topics, there was complete agreement. When confronted with atypical FNAC results, all participants agreed to send the patient to biopsy. When investigating the diagnostic weight the physicians place on certain patient characteristics, it was interesting to note that a family history of breast cancer had less weight than the mammographic result. Clinical impression also had a great deal of influence on the decision making, especially if the physician believed the lesion were cystic.

When discussing new technologies, the two that stand out presently in the diagnosis of breast cancer are stereotactic core biopsy and MRI. As with all new procedures, there are those who embrace them enthusiastically, and those who prefer to wait until the procedure is proven. But in general, the attitude among the participants in the groups was to watch and wait. Besides scientific reasons for this attitude, many cited the problem of insurance and referral. Many feared patients and the health system cannot finance all the new procedures, especially MRI, and that the greater need to refer patients to other specialists for different procedures would break down the continuity of the diagnostic process. In general, the main priority of diagnosis for all the physicians was to reduce the patient's anxiety about this process by providing the best and most complete diagnosis

as soon as possible. For this reason, many preferred FNAC since a preliminary result can be given in a few hours.

Before speaking to patients directly, we did a series of interviews with social workers who specialize in the area of counseling breast cancer patients or patients in the diagnostic process. These interviews and those with physicians served as background for the patient survey. The social workers were especially informative about patient attitudes towards the physicians who treat them, as well as the emotional and family conflicts which arise.

Having received IRB approval of our study protocol and our proposed patient survey (see **Appendix III**), we will undertake a series of patient interviews with several objectives in mind: 1) to investigate patient preferences, 2) to understand better patient concerns and beliefs about the diagnostic process, and 3) to elicit utilities useful for the decision analysis. The surveys will be administered in the radiology suites, gynecological clinic, and among support group members from two important local institutions. Whether these surveys will be self-administered or done by an interviewer will require further piloting work. A key consideration will be the development of time-tradeoff questions to assess utilities.

III. Conclusions

The foregoing narrative describes our progress to date at the close of the first project year. By doing so, we have provided a picture of our analytic strategy and our next steps. In project year 2, we intend to complete the entire structured literature review and meta-analysis focused on FNAC. We will also explore other statistical approaches to analyzing our database, such as examining the test characteristics as likelihood ratios[12] and considering the use of common ROC curves. We are now finishing the work on two physician focus groups and piloting the survey instrument for patients. The input from these streams of work will lay the foundation for our decision analysis on diagnostic strategies for evaluating breast abnormalities. This will culminate in a completed decision and cost-effectiveness analysis in project year 2.

REFERENCES

[1] Sutton SM, Doner LD. Insights into the Physician's Role in Mammography Utilization Among Older Women. *WHI* 1992; 2(4): 175-179.

[2] Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. *JAMA* 1993; 270(20): 2444-2450.

[3] Hayward RA, Shapiro MF, Freeman HE, Corey CR. Who Gets Screened for Cervical and Breast Cancer? Results from a New National Survey. *Arch Intern Med* 1988; 148: 1177-1181.

[4] Harris RP, Fletcher SW, Gonzalez JJ, et al. Mammography and Age: Are We Targeting the Wrong Women? A Community Survey of Women and Physicians. *Cancer* 1991; 67: 2010-2014.

[5] Light RJ, Pillemer DB. *Summing Up: The Science of Reviewing Research.* Cambridge, MA: Harvard University Press, 1984, pp. 63-65.

[6] Gulling, Pietro. Natural History of Cancer. *Cancer* 1977; 39: 2679-2730.

[7] Degenshein GA, Ceccarelli F, Bloom N, et al. Breast biopsies. Ratio of malignant to benign. *NY State J Med* 1976; 76: 1538.

[8] Spivey GH, Perry BW, Clark VA, Coulson AH, Coulson WF. Predicting the risk of cancer at the time of breast biopsy: Variation in the benign to malignant ratio. *The American Surgeon* 1982; 48(7): 326-332.

[9] Luft HS, Garnick DW, Mark DH, McPhee SJ. *Hospital Volume, Physician Volume, and Patient Outcomes: Assessing the Evidence.* Ann Arbor, MI: Health Administration Press, 1990.

[10] Cohen MB, Rodgers RPC, Hales MS, et al. Influence of Training and Experience in Fine-Needle Aspiration Biopsy of Breast. *Arch Pathol Lab Med* 1987; 111: 518-520.

[11] Berlin JA, Escarce JJ, Kinosian BP, Shea JA, Schwartz JS, Williams SV. Meta-Analysis of Diagnostic Test Characteristics (unpublished manuscript), p. 15.

[12] Giard RWM, Hermans J. Interpretation of Diagnostic Cytology with Likelihood Ratios. *Arch Pathol Lab Med* 1990; 114: 852-854.

APPENDIX I

Meta-Analysis Review Form for Diagnostic Test Articles

Article ID number: **First Author:Last name, First initial.**
Journal and Date of Publication:

Reviewer's name: 1=Jesse Berlin 2=Stephen Clyman 3=Suzanne Fletcher 4=Kathy Hirata 5=Anthony So 6=John Wong 7=Joseph Yi 8=Gwen Barretto 9=Vincenza Snow

Diagnostic test:

1. Patient self examination
2. Breast clinical examination
3. Mammography
4. Fine-needle aspiration cytology (FNAC)
5. Core biopsy
6. Tru Cut biopsy
7. Excisional breast biopsy
8. Ultrasonography
9. Thermography
10. MRI

Equipment description:

Tissue prep: (for biopsy)

Localizations: (for biopsy)

1. Palpable
2. Stereotactic
3. Ultrasonography
4. Mammography
5. NM
6. None of the above (specify):

Does study evaluate: 1. a single test
2. multiple tests

Sequences of diagnostic tests used in study:

Reject
Accept

Verification Bias: Yes / No

Reason for Rejection:

1. Not relevant
2. N < 100 (number of tests)
3. No original data (e.g., review article)
4. Absent gold or reference standard
5. Special patient population -- Please specify:
6. Verification bias
7. Procedural variation
8. Special subset
9. Other

Sample Size:

- N -- cases reviewed
- N -- tests included in study
  - N excluded
  - N dropped out
  - N inadequate/ nondiagnostic
- N -- subjects
  - N excluded
  - N dropped out
  - Significant prestudy exclusions

Describe exclusions (Before study, After study, Draw tree):

Types of breast cancer

0. All types
1. Ductal
2. Lobular
3. Ductal CIS
4. Lobular CIS
5. Papillary
6. Colloid
7. Medullary
8. Paget's
9. Apocrine
10. Tubular

1. Palpable
2. Nonpalpable
3. Both
4. NM

Patient Descriptions:

Sex Female _____ Male _____ Not mentioned

Age Information: 1. present 2. absent Table? yes no

Measure used to report age: 1. Mean 2. Median 3. Range 4. None

	Overall	Subgroup 1	Subgroup 2	Subgroup 3
Measure				
Subgroup characterized				

Diagnostic work-up or tests at baseline entry

Defined yes no inferred

Tests at baseline:

1. Patient self examination
2. Breast clinical examination
3. Mammography
4. Fine-needle aspiration cytology (FNAC)
5. Core biopsy
6. Tru cut biopsy
7. Excisional breast biopsy
8. Ultrasonography
9. Thermography
10. MRI

Were observers blinded to result of index and gold standard/reference test? Yes/No/NM
Gold standard for positive test

Type of Biopsy

Biopsy result **All cases / Selected cases / None**
n =
Clinical follow-up **All cases / Selected cases / None**
n =

Clinical follow-up Data:
1. Yes 2. No 3. NM (Not mentioned)

Follow-up range				
LOW	**days**	**weeks**	**months**	**NM**
HIGH	**days**	**weeks**	**months**	**NM**
Mean/Median	**days**	**weeks**	**months**	**NM**

Lost to follow-up: **1. Mentioned 2. Not mentioned**
Characterize those lost to follow-up: _____

_____

_____

_____

Spectrum of disease: **Mentioned Not mentioned**
Characteristics of cancer detected (Tumor/Metastatic/Nodes, size of tumors):

Test characteristics:
No. of times procedure is performed for each test (e.g., aspirations): **or NM**

Population or subpopulation

Diagnostic categories **Construct n x n table**

		Gold	
Dx test		D+	D-
T+	Malignant		
T+	Suspicious		
T+	Atypical		
T-	Benign		
	Inadequate		

	D+	D-
T+		
T-		

Sensitivity =TP/TP+FN

Specificity =TN/TN+FP

Prevalence = # D+/total # cases

Positive predictive value TP/TP+FP

Negative predictive value TN/TN+FN

Provider(s): No. and Description:

_____

Operator(s) description:

a. Aspirators:_____

a1). No. of aspirators: 1. One 2. More than one 3. NM

b. Cytopathologists (or other) interpreting the FNA specimens:

_____

_____

b1). No. of people reading FNA specimens: 1. One 2. More than one 3. NM

c. Inter-rater reliability measures Yes / No

If yes, describe: _____

d. Training or experience noted Yes / No

Describe:_____

d1). Did False Positives, Inadequates, etc. occur early in series? Describe:

_____

_____

Facility where test was performed (name, location, and describe):

_____

_____

Duration/Time Period of Patient Recruitment into the study:

_____

Complications:

No / Yes / NM

Table? No / Yes

Type of complication:

Total	Group 1	Group 2	Group 3	Group 4
#				
%				

Type of complication:

Total	Group 1	Group 2	Group 3	Group 4
#				
%				

Type of complication:

Total	Group 1	Group 2	Group 3	Group 4
#				
%				

Mortality?

1. No 2. Yes 3. NM Table? No/ Yes

Total	Group 1	Group 2	Group 3	Group 4
#				
%				

Other Outcomes?

1. No 2. Yes 3. NM

Please specify type and frequency:

Double counting under complications?

1. NA (no complications)
2. No
3. Yes
4. Not sure
5. NM

Design:

1. Case Series
2. Consecutive series
3. Cohort
4. Randomized
5. Case-control
6. Other
7. NM

Data collection approach:

1. Chart review
2. Questionnaire
3. Claims analysis
4. Other -- please specify:

Time Frame:

1. Prospective
2. Retrospective
3. Other
4. NM

Comments / Concerns:

No/ Yes

If yes, please specify: _____

_____

_____

_____

_____

_____

_____

_____

Backpage information

Reject
Accept
Hold for further discussion

Does the information from this article offer data redundant with another article? Yes / No

If so, what was the other article citation?

Appendix II
Bibliography of Articles Accepted for Meta-Analysis

1. M. J. Allen, et al., *Br J Surg* **81**, 543-5 (1994).106
2. H. T. Aretz, M. L. Silverman, J. L. Kolodziejski, B. R. Witherspoon, *Postgrad Med* **75**, 49-52 (1984).1267
3. M. Bibbo, F. P. Zuspan, *Am J Obstet Gynecol* **122**, 525-6 (1975).1405
4. G. S. Daum, T. S. Kline, R. L. Artymyshyn, H. S. Neal, *Cancer* **67**, 2150-2 (1991).600
5. J. E. Devitt, R. H. Curry, *Can J Surg* **20**, 450-1 (1977).1471
6. K. Dowlatshahi, P. M. Jokich, L. F. Kluskens, R. Patel, S. G. Economou, *Surg Gynecol Obstet* **172**, 121-4 (1991).487
7. L. L. Fajardo, J. R. Davis, J. L. Wiens, D. C. Trego, *Ajr Am J Roentgenol* **155**, 977-81 (1990).632
8. W. J. Frable, *Cancer* **53**, 671-6 (1984).1273
9. T. I. Gardecki, B. M. Hogbin, D. H. Melcher, R. S. Smith, *Lancet* **2**, 790-2 (1980).1189
10. R. K. Gupta, A. G. McHutchison, J. S. Simpson, C. S. Dowle, *Acta Cytol* **35**, 703-9 (1991).468
11. S. Hammond, S. Keyhani-Rofagha, R. V. O'Toole, *Acta Cytol* **31**, 276-80 (1987).1046
12. C. Hermansen, et al., *Acta Chir Scand* **150**, 625-8 (1984).1240
13. D. M. Ingram, G. F. Sterrett, H. J. Sheiner, K. B. Shilkin, *Med J Aust* **2**, 170-3 (1983).1290
14. M. P. Kahky, V. R. Rone, D. L. Duncan, A. Cruz Jr., H. 3. Gaskill, *Am J Surg* **156**, 450-2 (1988).920
15. M. Kaufman, D. Bider, D. Weissberg, *Am Surg* **49**, 558-9 (1983).1279
16. T. S. Kline, H. S. Neal, *Acta Cytol* **20**, 324-7 (1976).1490
17. T. S. Kline, H. S. Neal, *Jama* **239**, 36-9 (1978).1424
18. A. P. Koivuniemi, *Ann Clin Res* **8**, 272-83 (1976).1502
19. G. Kreuzer, E. Boquoi, *Acta Cytol* **20**, 319-23 (1976).1491
20. J. Lamb, T. J. Anderson, M. J. Dixon, P. A. Levack, *J Clin Pathol* **40**, 705-9 (1987).1083
21. V. K. Langmuir, S. F. Cramer, M. E. Hood, *Acta Cytol* **33**, 93-8 (1989).845
22. L. H. Manheimer, A. M. Rywlin, *South Med J* **70**, 923-5 (1977).1475
23. I. Nyirjesy, F. S. Billingsley, *Obstet Gynecol* **79**, 699-702 (1992).427
24. S. R. Orell, *Med J Aust* **2**, 163-8 (1976).1505
25. D. Rotten, J. M. Levaillant, H. Leridon, A. Letessier, M. Sandres, *Eur J Obstet Gynecol Reprod Biol* **49**, 175-86 (1993).233
26. N. A. Sharifah, A. R. Hayati, F. Meah, Z. al-Rashid, *Malays J Pathol* **15**, 53-8 (1993).191
27. R. G. Somers, et al., *Arch Surg* **120**, 673-7 (1985).1175
28. H. T. Strawbridge, A. A. Bassett, I. Foldes, *Surg Gynecol Obstet* **152**, 1-7 (1981).1370
29. J. M. Thomas, et al., *Br Med J* **2**, 1139-41 (1978).1453
30. R. M. Ulanow, L. Galblum, J. W. Canter, *Am J Surg* **148**, 653-7 (1984).1247
31. K. Verma, K. Kapila, *Indian J Med Res* **90**, 135-9 (1989).819
32. A. Zajdela, N. A. Ghossein, J. P. Pilleron, A. Ennuyer, *Cancer* **35**, 499-506 (1975).1537
33. G. D. Stavric, D. T. Tevcev, D. R. Kaftandjiev, J. J. Novak, *Acta Cytol* **17**, 188-90 (1973).1555
34. S. I. Hajdu, M. R. Melamed, *Am J Clin Pathol* **59**, 350-6 (1973).1574
35. M. Bibbo, et al., *Acta Cytol* **32**, 193-201 (1988).879
36. C. Smith, J. Butler, C. Cobb, D. State, *Surgery* **103**, 178-83 (1988).994
37. F. A. Sheikh, G. H. Tinkoff, T. S. Kline, H. S. Neal, *Am J Surg* **154**, 470-4 (1987).1070
38. S. Masood, E. R. Frykberg, G. L. McLellan, J. B. Bullard, *South Med J* **84**, 857-61 (1991).566
39. E. Azavedo, G. Svane, G. Auer, *Lancet* **1**, 1033-6 (1989).735
40. J. A. Zuk, G. Maudsley, H. D. Zakhour, *J Clin Pathol* **42**, 906-11 (1989).767
41. S. Ciatto, M. Rosselli Del Turco, P. Bravetti, *Radiology* **173**, 57-9 (1989).768
42. L. J. Layfield, B. J. Glasgow, H. Cramer, *Pathol Annu* **2**, 23-62 (1989).772
43. D. C. Knight, D. M. Lowell, A. Heimann, E. Dunn, *Surg Gynecol Obstet* **163**, 415-20 (1986).1112
44. C. N. Griffith, W. H. Kern, W. P. Mikkelsen, *Surg Gynecol Obstet* **162**, 142-4 (1986).1134

45. L. Deschenes, et al., *Can J Surg* **21**, 417-9 (1978).1451
46. S. H. Parker, et al., *Radiology* **193**, 359-64 (1994).27
47. R. H. Janes, M. S. Bouton, *Am J Surg* **168**, 533-6 (1994).24
48. J. S. Caines, M. D. McPhee, G. P. Konok, B. A. Wright, *Ajr Am J Roentgenol* **163**, 317-21 (1994).90
49. J. J. Gisvold, et al., *Ajr Am J Roentgenol* **162**, 815-20 (1994).159
50. F. Burbank, K. Kaye, J. Belville, J. Ekuan, M. Blumenfeld, *Radiology* **191**, 165-71 (1994).162
51. E. L. Elvecrog, M. C. Lechner, M. T. Nelson, *Radiology* **188**, 453-5 (1993).299
52. S. H. Parker, et al., *Radiology* **180**, 403-7 (1991).476
53. D. P. Watson, M. McGuire, F. Nicholson, H. F. Given, *Surg Gynecol Obstet* **165**, 435-41 (1987).885
54. S. A. Dundas, P. R. Sanderson, H. Matta, A. J. Shorthouse, *Acta Cytol* **32**, 202-6 (1988).878
55. L. A. Brown, S. B. Coghill, *Cytopathology* **2**, 67-74 (1991).547
56. H. J. Wanebo, P. S. Feldman, M. C. Wilhelm, J. L. Covell, R. L. Binns, *Ann Surg* **199**, 569-79 (1984).1261
57. J. F. Silverman, D. R. Lannin, K. O'Brien, H. T. Norris, *Acta Cytol* **31**, 731-6 (1987).1061
58. P. S. Cheung, K. W. Yan, T. T. Alagaratnam, *Aust N Z J Surg* **57**, 615-20 (1987).1064
59. B. Adye, P. C. Jolly, D. E. Bauermeister, *Arch Surg* **123**, 37-9 (1988).996
60. D. M. Hansell, J. C. Cooke, C. A. Parsons, *Clin Radiol* **39**, 150-3 (1988).935
61. G. Svane, *Recent Results Cancer Res* **105**, 95-6 (1987).1016
62. B. D. Fornage, M. J. Faroux, A. Simatos, *Radiology* **162**, 409-14 (1987).1022
63. G. Hiller, M. Harney, S. Legg, S. A. Hart, *Aust N Z J Surg* **57**, 239-42 (1987).1037
64. M. R. Sarfati, et al., *Am J Surg* **168**, 529-31 (1994).25
65. D. J. Oliver, J. R. Frayne, G. Sterrett, *Aust N Z J Surg* **62**, 463-7 (1992).417
66. M. A. Helvie, et al., *Radiolcgy* **174**, 657-61 (1990).724
67. C. Hermansen, et al., *Cancer* **60**, 1866-71 (1987).1078
68. T. T. Casey, et al., *Am J Surg* **163**, 305-11 (1992).448
69. A. I. Cook, K. B. Robinson, *J R Coll Surg Edinb* **36**, 289-92 (1991).507
70. M. Lofgren, I. Andersson, K. Lindholm, *Recent Results Cancer Res* **119**, 100-4 (1990).670
71. J. S. Mitnick, et al., *Arch Surg* **126**, 1137-40 (1991).472
72. G. Saunders, Y. Lakra, J. Libcke, *Surg Gynecol Obstet* **172**, 437-40 (1991).481
73. N. Sneige, *Diagn Cytopathol* **9**, 106-12 (1993).333
74. G. Sterrett, D. Oliver, J. Frayne, D. Ingram, H. Sheiner, *Pathology* **23**, 302-10 (1991).466
75. H. S. Vine, J. E. Jacobs, E. B. Cronin, W. E. Clark, *Conn Med* **58**, 67-70 (1994).99
76. K. Dowlatshahi, et al., *Radiology* **170**, 427-33 (1989).788
77. A. J. Howat, H. F. Stringfellow, W. A. Briggs, C. M. Nicholson, *Acta Cytol* **38**, 939-44 (1994).10
78. S. Negri, F. Bonetti, A. Capitanio, M. Bonzanini, *Diagn Cytopathol* **11**, 4-8 (1994).34
79. Z. Kaufman, et al., *J Surg Oncol* **56**, 254-7 (1994).78
80. M. J. Kerin, et al., *Br J Surg* **81**, 221-2 (1994).144
81. S. E. Iles, J. H. MacGregor, A. J. Bodurtha, A. I. Bernardo, C. Daniels, *Can Assoc Radiol J* **45**, 28-34 (1994).170
82. P. B. Gordon, S. L. Goldenberg, N. H. Chan, *Radiology* **189**, 573-80 (1993).231
83. M. J. Costa, T. Tadros, G. Hilton, G. Birdsong, *Acta Cytol* **37**, 461-71 (1993).297
84. S. Ciatto, S. Catarzi, D. Morrone, M. R. Del Turco, *Radiology* **188**, 195-8 (1993).312
85. S. Rotstein, B. Nilsson, G. Svane, E. Gustavson-Kadaka, *Acta Oncol* **31**, 393-7 (1992).405
86. K. Dowlatshahi, M. L. Yaremko, L. F. Kluskens, P. M. Jokich, *Radiology* **181**, 745-50 (1991).524
87. G. R. Nicastri, W. P. Reed, B. R. Dziura, *Surg Gynecol Obstet* **172**, 457-60 (1991).582
88. P. G. Horgan, et al., *Surg Gynecol Obstet* **172**, 290-2 (1991).598

Appendix II
Bibliography of Articles Accepted for Meta-Analysis

89. S. Masood, et al., *Cancer* **66**, 1480-7 (1990).616
90. H. A. Gelabert, J. G. Hsiu, J. T. Mullen, A. H. Jaffe, N. A. D'Amato, *Am Surg* **56**, 263-7 (1990).619
91. E. J. Wilkinson, C. M. Schuettke, C. M. Ferrier, D. A. Franzini, K. I. Bland, *Acta Cytol* **33**, 613-9 (1989).812
92. W. H. Wolberg, M. A. Tanner, W. Y. Loh, *Arch Surg* **124**, 814-8 (1989).824
93. R. W. Painter, W. 2. Clark, P. J. Deckers, *Am J Surg* **155**, 387-90 (1988).991
94. N. Sneige, B. D. Fornage, G. Saleh, *Am J Clin Pathol* **102**, 98-101 (1994).89
95. M. Lofgren, I. Andersson, K. Lindholm, *Ajr Am J Roentgenol* **154**, 1191-5 (1990).634
96. J. J. Patel, P. C. Gartell, P. B. Guyer, A. Herbert, I. Taylor, *J R Soc Med* **81**, 10-2 (1988).939
97. R. J. Zarbo, P. J. Howanitz, P. Bachner, *Arch Pathol Lab Med* **115**, 743-50 (1991).564
98. J. Frisell, G. Eklund, R. Nilsson, L. Hellstrom, A. Somell, *Br J Surg* **76**, 840-3 (1989).817
99. M. F. Vazquez, J. S. Mitnick, P. Pressman, M. N. Harris, D. F. Roses, *J Am Coll Surg* **178**, 17-23 (1994).146
100. F. I. Atamdede, J. H. Isaacs, *Gynecol Oncol* **50**, 159-63 (1993).282

APPENDIX III

PATIENT SURVEY QUESTIONS

Directions: (Some instructions for filling out the form)

1. Sex: O Female
   O Male

2. Age: OO

2. Who sent you for a Mammogram?
   - O Ob/Gyn
   - O Family Practitioner
   - O Internist
   - O Other

3. How many times have you had a mammogram?
   - O One
   - O Two or Three
   - O More than three

4. When was your last mammogram?

   OO / OO / OO
   m / d / y

5. Are you receiving a mammogram for (check all that apply):
   - O Screening (Routine check)
   - O Breast Lump
   - O Breast Symptoms
   - O Follow-Up

6. From a scale of 1 to 10, how would you rate your anxiety about the result of the mammogram? (Circle one)

   1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 10
   None Moderate Extremely

7. If for follow-up, what is your current diagnosis

   _____

8. Have you had any kind of diagnostic procedure done before your mammogram? (check all that apply)
   - O UltraSound
   - O Fine Needle Aspiration Cytology
   - O Thermography
   - O Others (please specify)

9. If for a lump, who detected it?
   - O Myself
   - O My Physician
   - O Other

10. Do you do a Breast Self Exam on yourself regularly?
    - O Yes
    - O No

    If Yes, how often? _____

11. Have you had any kind of breast problems before now?
    - O Yes
    - O No

12. If yes, please note what the diagnosis was

    _____

13. Which of the following applies to you?
    - O Family member had/has breast cancer
    - O Family member has been treated for breast cancer
    - O Friend has/had breast cancer
    - O Friend has been treated for cancer

14. Have you ever had a mammogram that said that you might have cancer but you did not?
    - O Yes
    - O No

15. What procedures if any did you undergo to find out that there was no cancer ?
    - O Ultrasound
    - O Biopsy
    - O Fine Needle Aspiration
    - O Other(please specify)

16. Has that experience ever caused you not to want to continue with screening mammograms?
    - O Yes
    - O No

17. Do you believe all lumps must be biopsied?
    - O Yes
    - O No

18. Do you believe all breast lumps must be surgically removed?
    - O Yes
    - O No

19. Can a needle or surgical biopsy cause a cancer to spread?
    - O Yes
    - O No

20. Can mammography detect all cancers?
    - O Yes
    - O No

    Can it substitute for breast self exam?
    - O Yes
    - O No

---

21. Would you prefer
    - O a needle biopsy with 80-90% certainty of diagnosis and no scar
    - O Or an excisional biopsy with 100% certainty of diagnosis,and some scarring

22. Would you prefer
    - O a needle biopsy with almost 100% certainty of diagnosis, no scarring, but the possibility you will have to undergo excisional surgery anyway
    - O Or an excisional surgical biopsy with scarring

23. Would you prefer
    - O a mammogram that detects 85-90% of cancers but is inexpensive and covered by insurance
    - O Or a much more expensive MRI which is not covered by insurance , but that detects 95% or more of cancers

24. If you had a needle biopsy and the result was uncertain, would you prefer
    - O A repeat needle biopsy since it leaves no scar, but the result could again be uncertain
    - O An excisional biopsy and scarring but have the certainty of a diagnosis

25. If you had a benign(normal) result on a biopsy ,would you return for follow-up visits and tests 3 or 6 months later if you have absolutely no symptoms ?
    - O Yes
    - O No
    - O Maybe

26. If you received an uncertain result on a biopsy, would you return for follow-up visits and tests 3 or 6 months later if you have absolutely no symptoms ?
    - O Yes
    - O No
    - O Maybe

27. If you had to chose, which do you prefer...
    - O Conservative surgery but the risk of the lump or cancer returning
    - O Mastectomy but the knowledge that the tumor will not return

Appendix IV: Charts

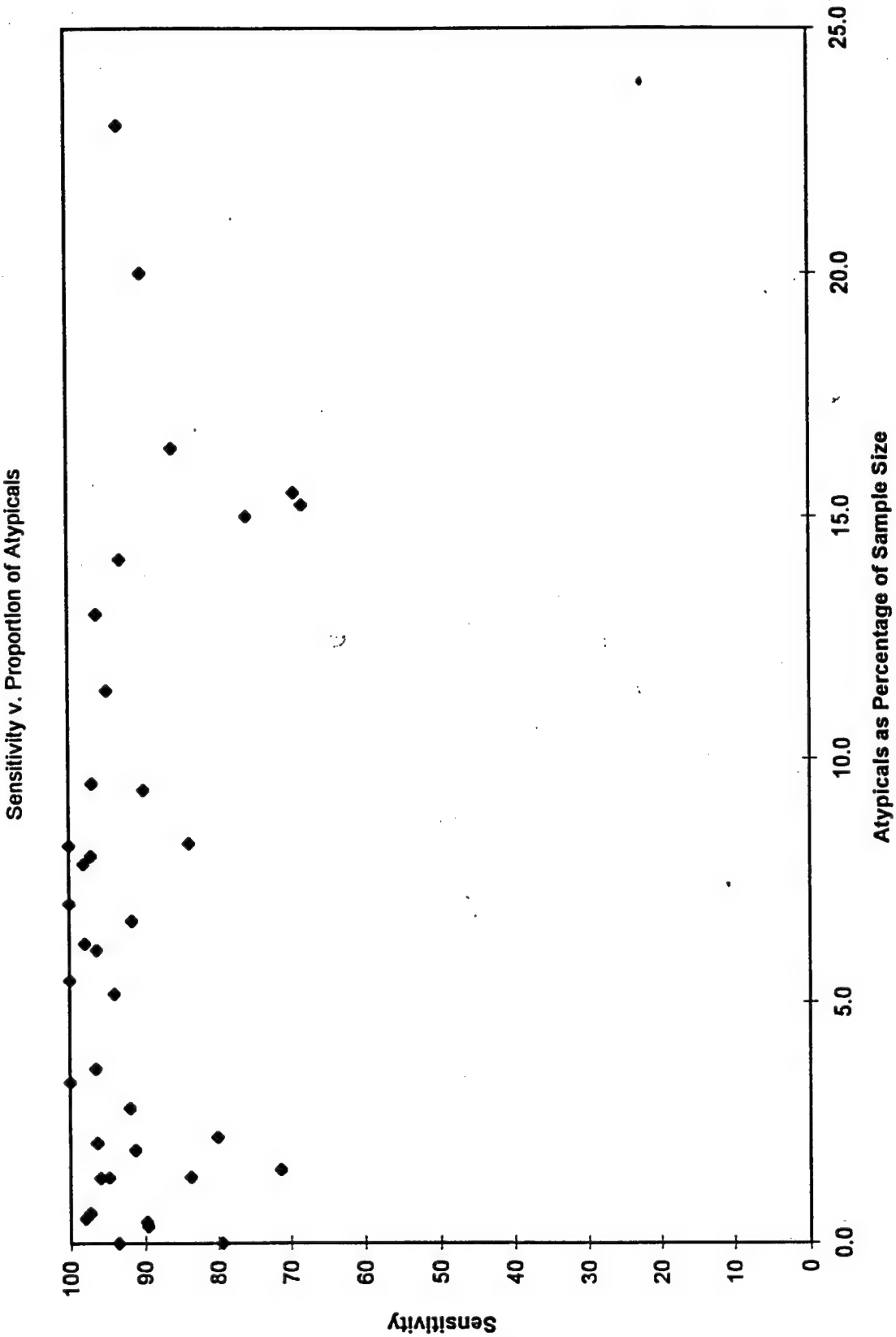


Chart 1
Sensitivity v. Proportion of Atypicals
Sensitivity
100
90
80
70
60
50
40
30
20
10
0
0.0
5.0
10.0
15.0
20.0
25.0
Atypicals as Percentage of Sample Size

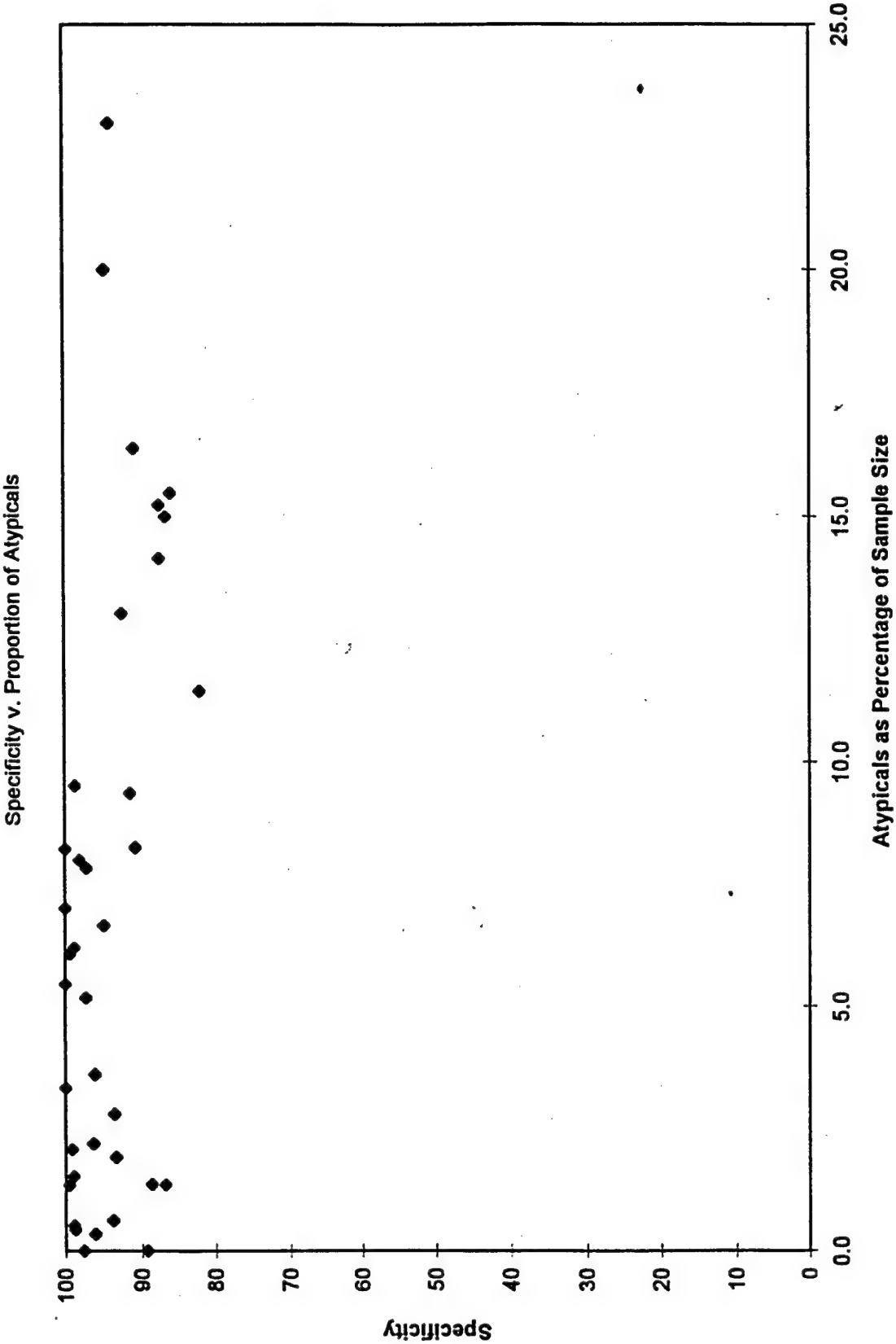


Chart 2
Specificity v. Proportion of Atypicals
Specificity
100
90
80
70
60
50
40
30
20
10
0
0.0
5.0
10.0
15.0
20.0
25.0
Atypicals as Percentage of Sample Size

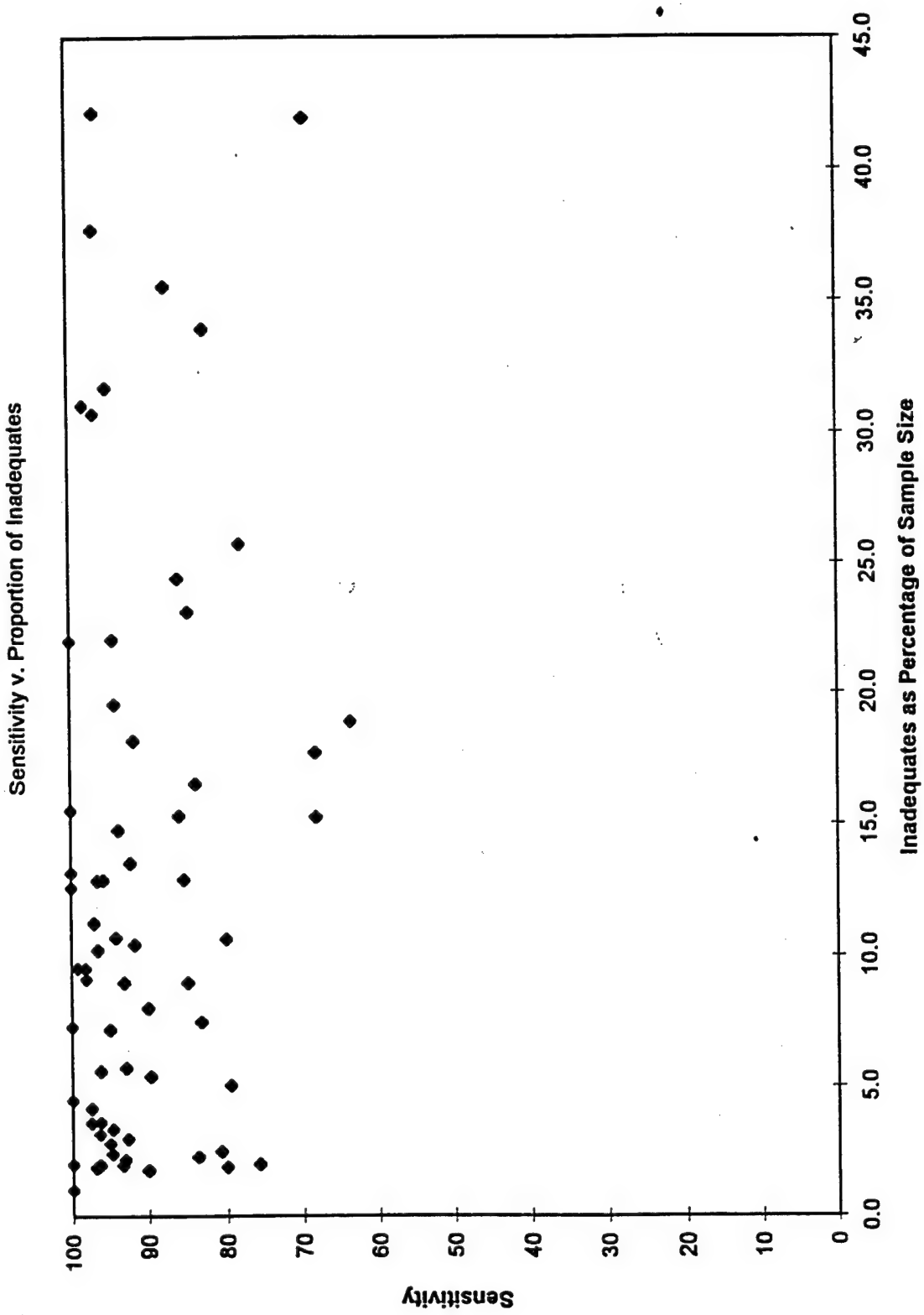


Chart 3
Sensitivity v. Proportion of Inadequates
Sensitivity
100
90
80
70
60
50
40
30
20
10
0
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
Inadequates as Percentage of Sample Size

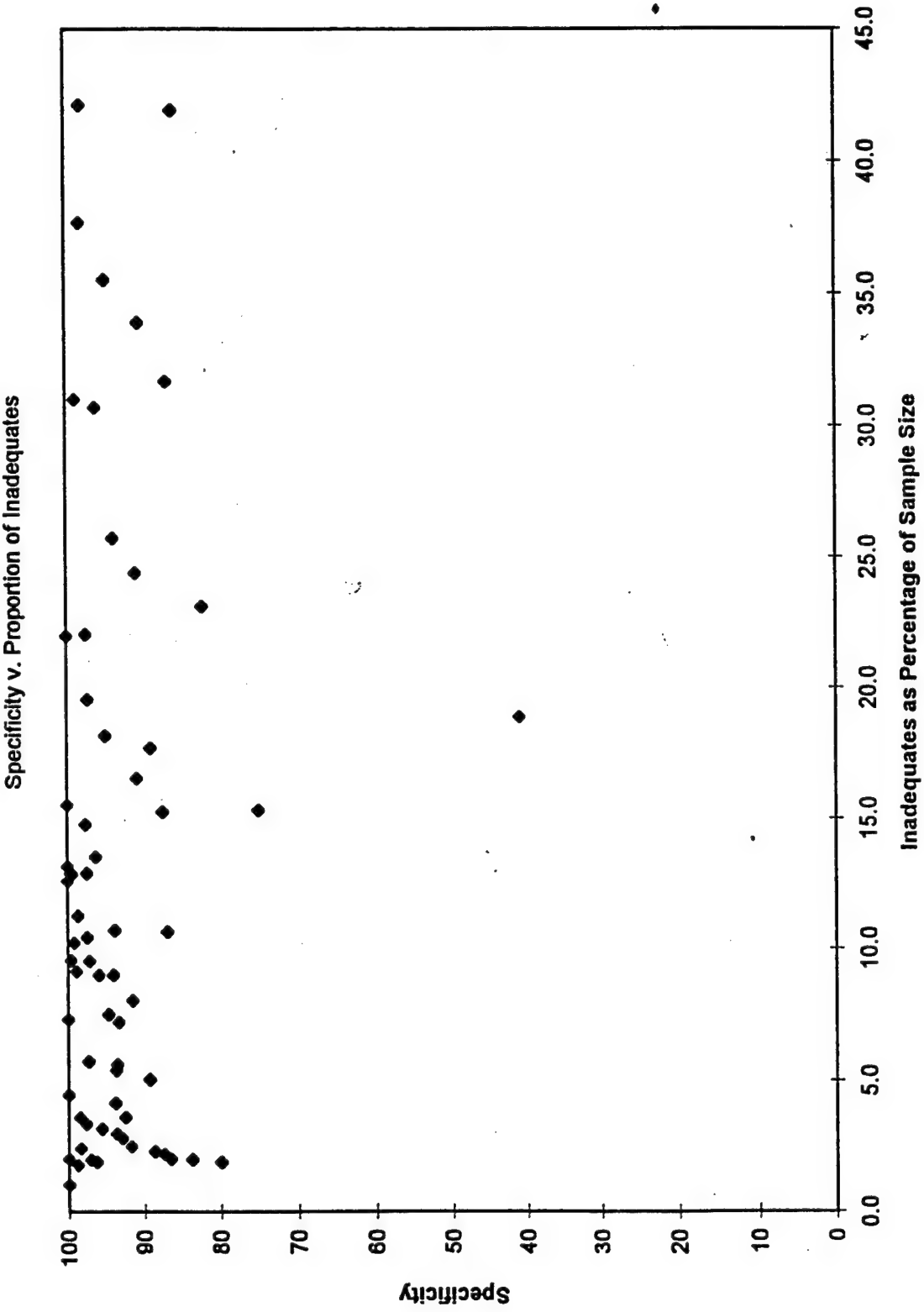
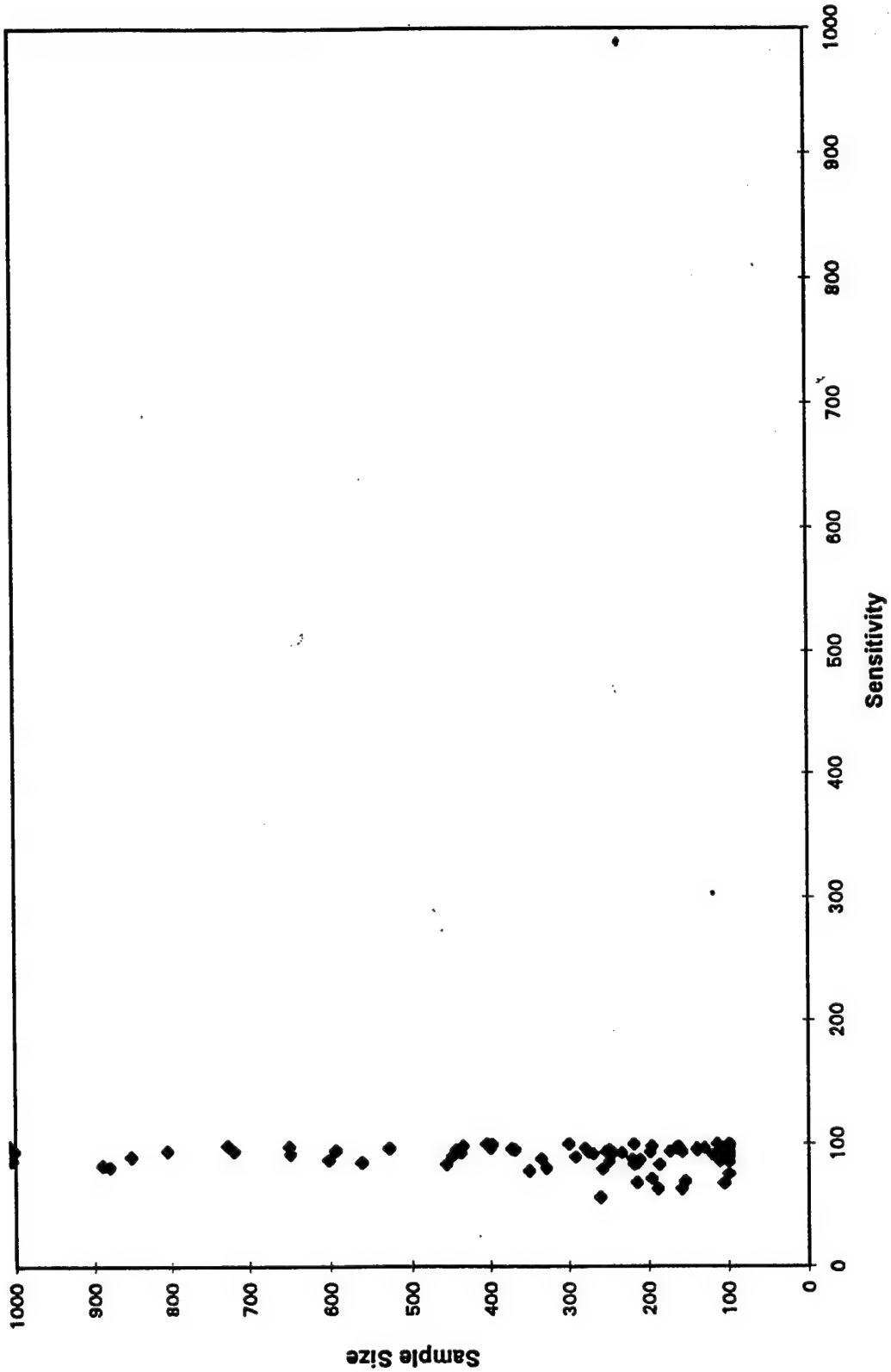


Chart 4
Specificity v. Proportion of Inadequates
Specificity
100
90
80
70
60
50
40
30
20
10
0
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
Inadequates as Percentage of Sample Size

Chart 5
Relationship between Sample Size and Sensitivity



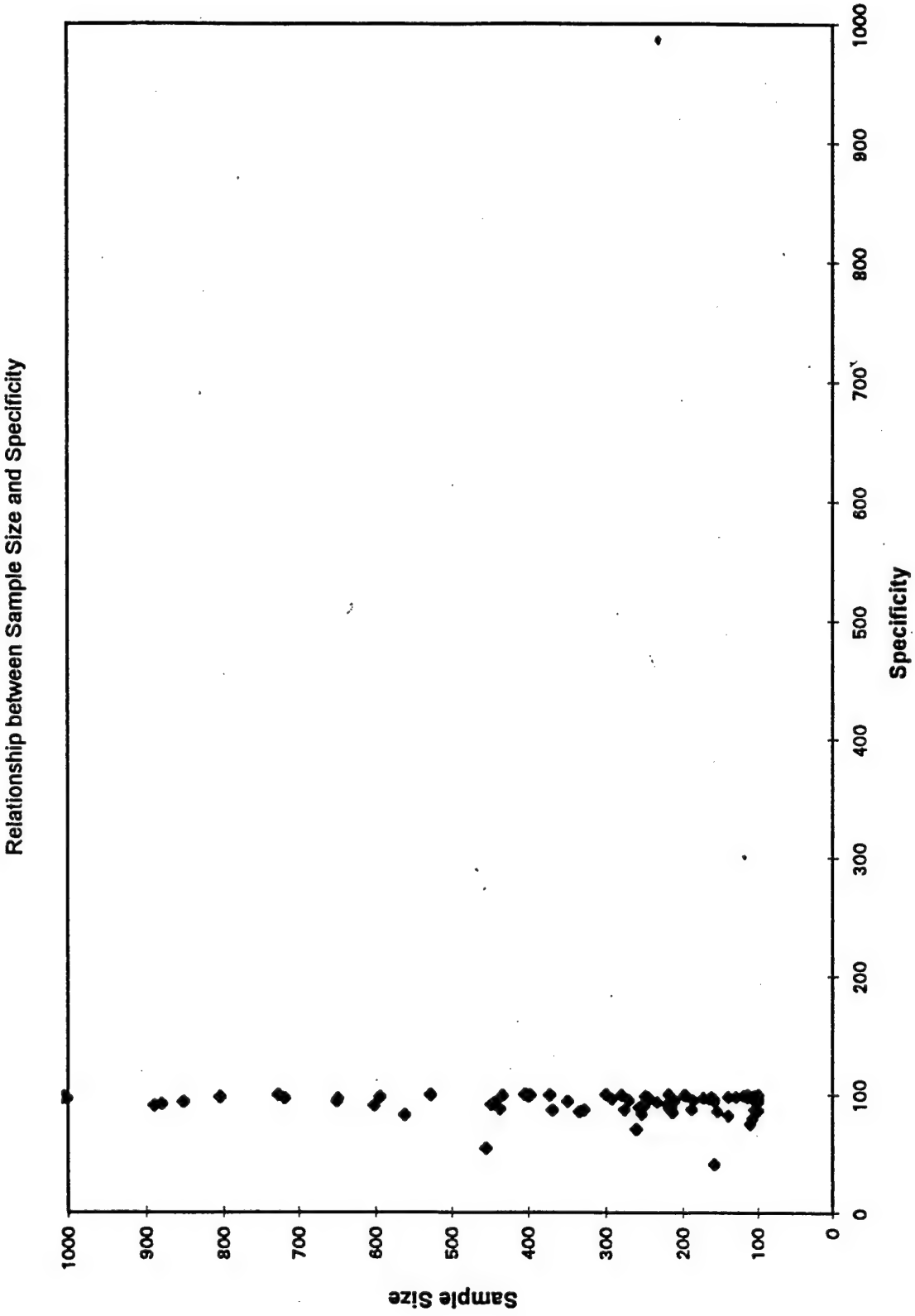


Chart 6
Relationship between Sample Size and Specificity
Sample Size
Specificity
1000
900
800
700
600
500
400
300
200
100
0
0
100
200
300
400
500
600
700
800
900
1000

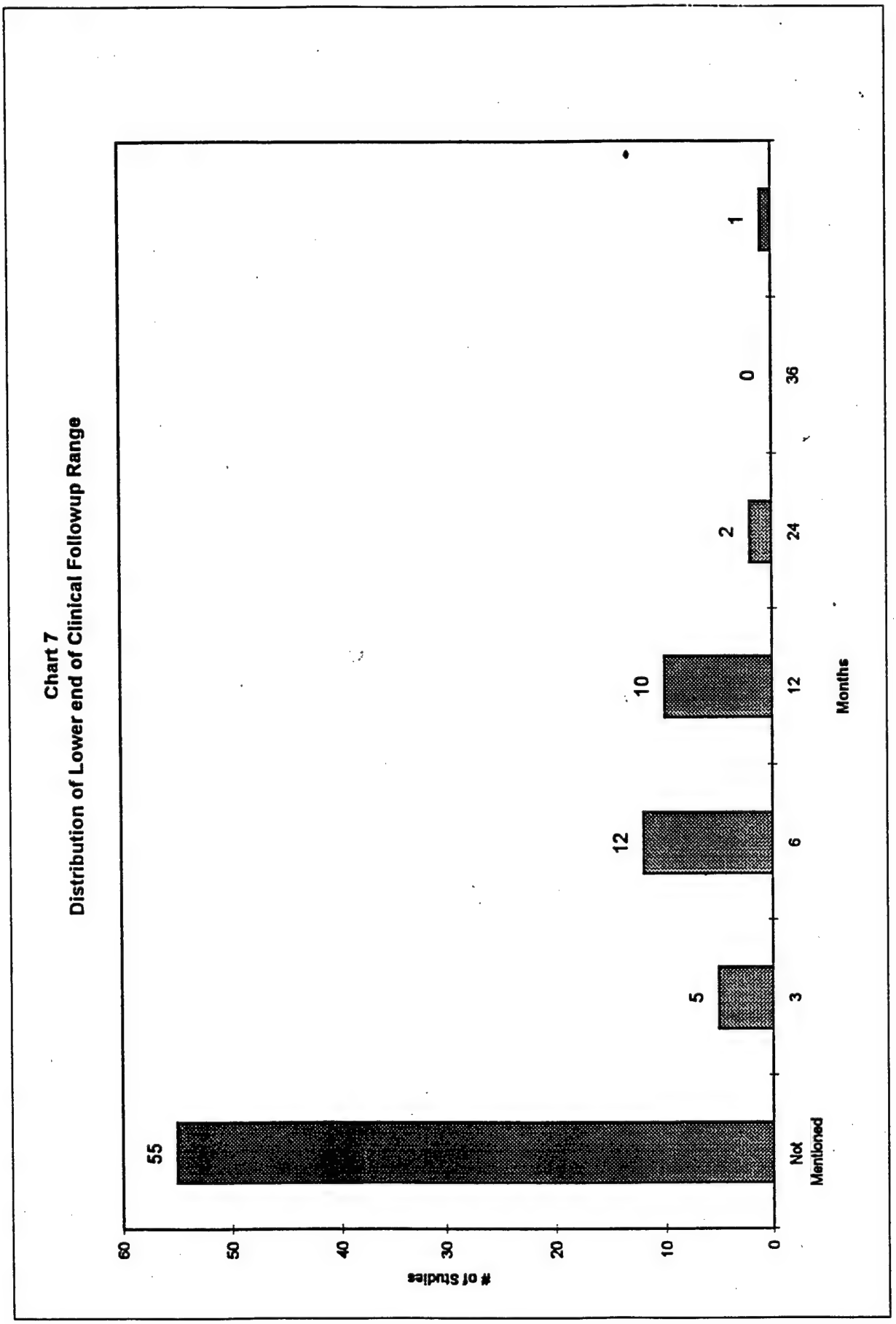
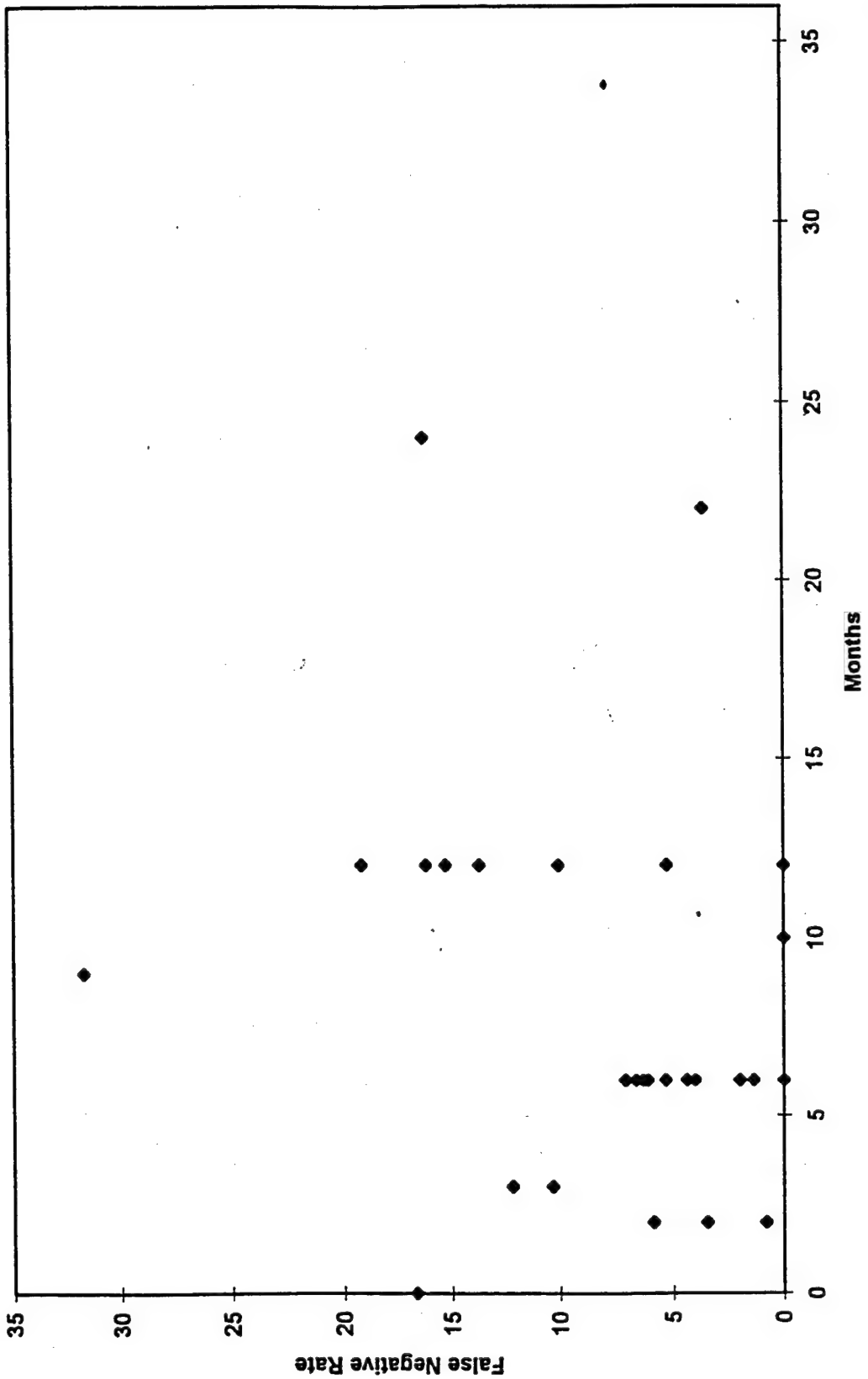


Chart 7
Distribution of Lower end of Clinical Followup Range
# of Studies
60
50
40
30
20
10
0
55
5
12
10
2
0
1
Not Mentioned
3
6
12
24
36
Months

Chart 8
Relationship between False Negative Rate and Lower Range of Clinical Followup



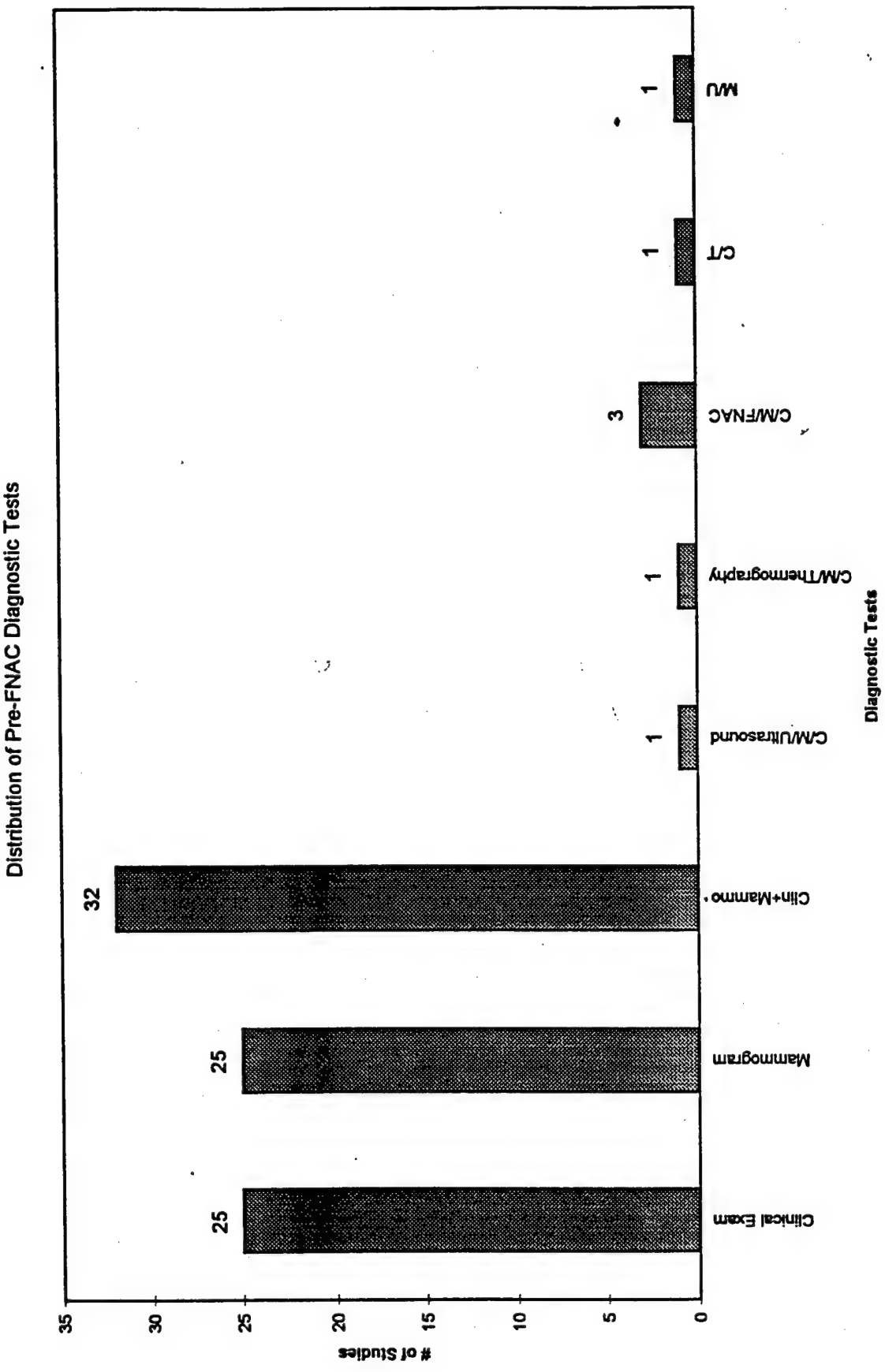


Chart 9
Distribution of Pre-FNAC Diagnostic Tests
# of Studies
0
5
10
15
20
25
30
35
Clinical Exam
25
Mammogram
25
Clin+Mammo
32
C/M/Ultrasound
1
C/M/Thermography
1
C/M/FNAC
3
C/T
1
M/U
1
Diagnostic Tests

Chart 10
Sensitivity v. Benign:Malignant Ratio

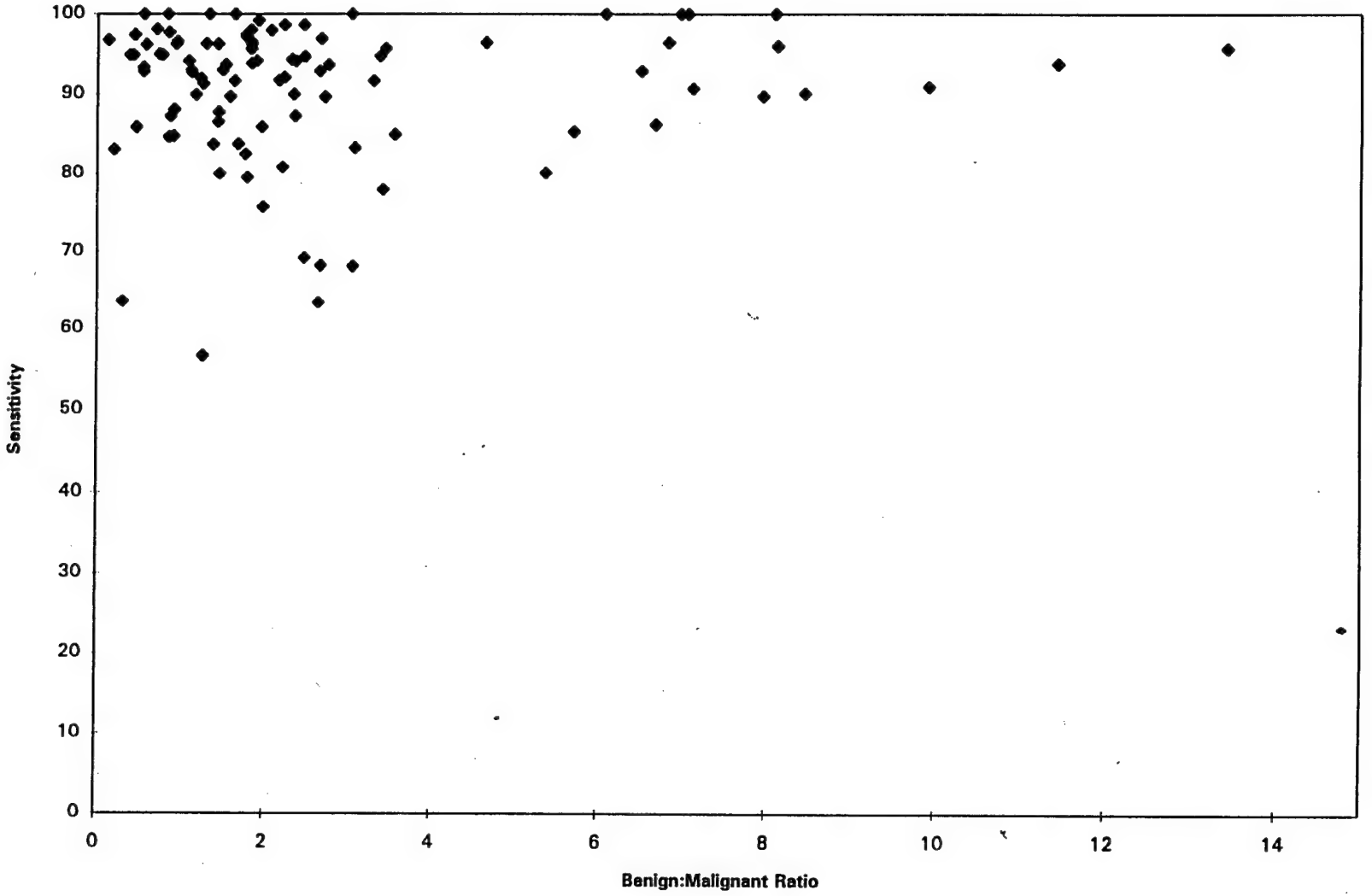


Chart 11
Specificity v. Benign:Malignant Ratio

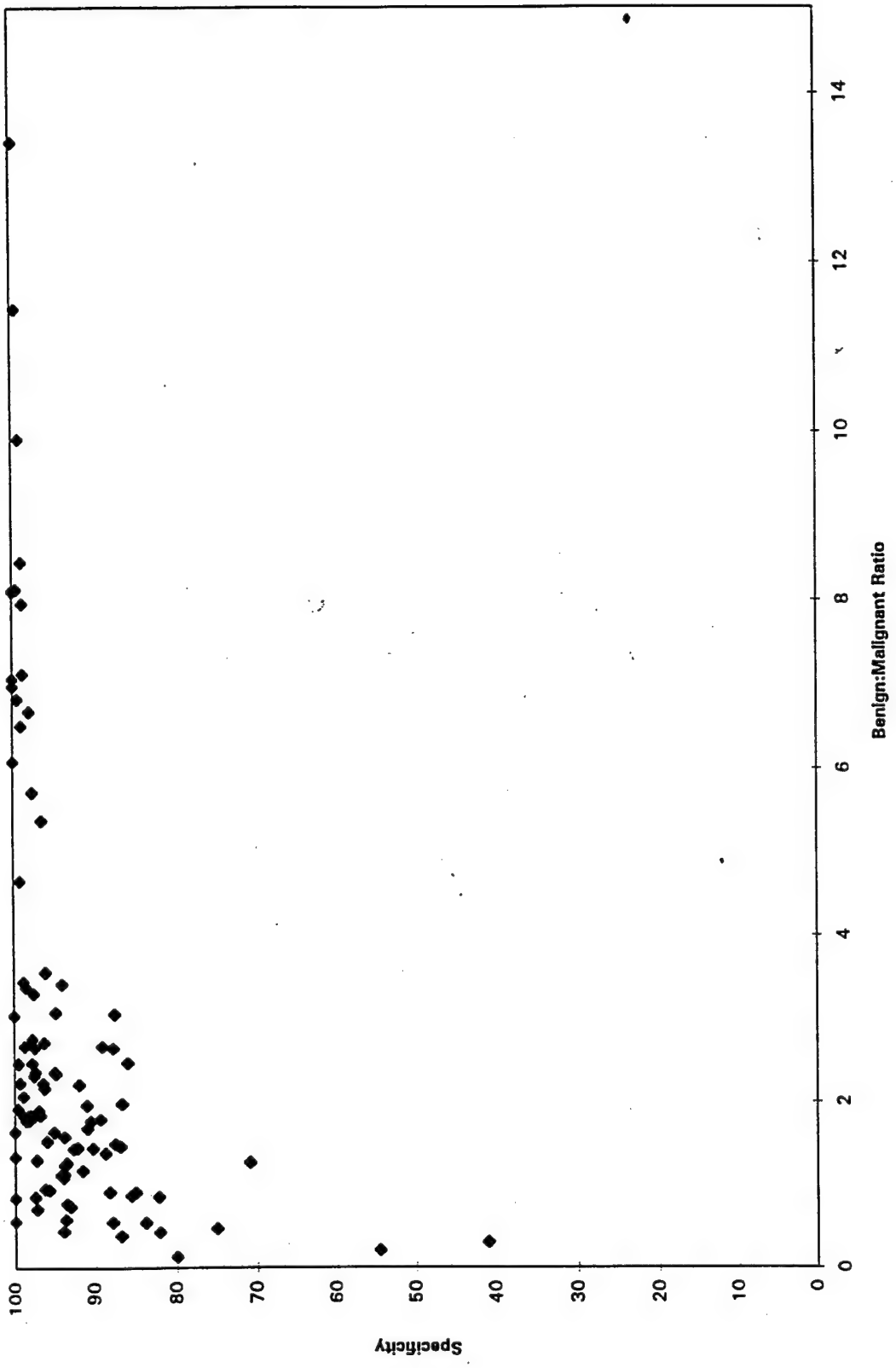


Chart 12

Sensitivity v. Volume per Operator per Unit of Time

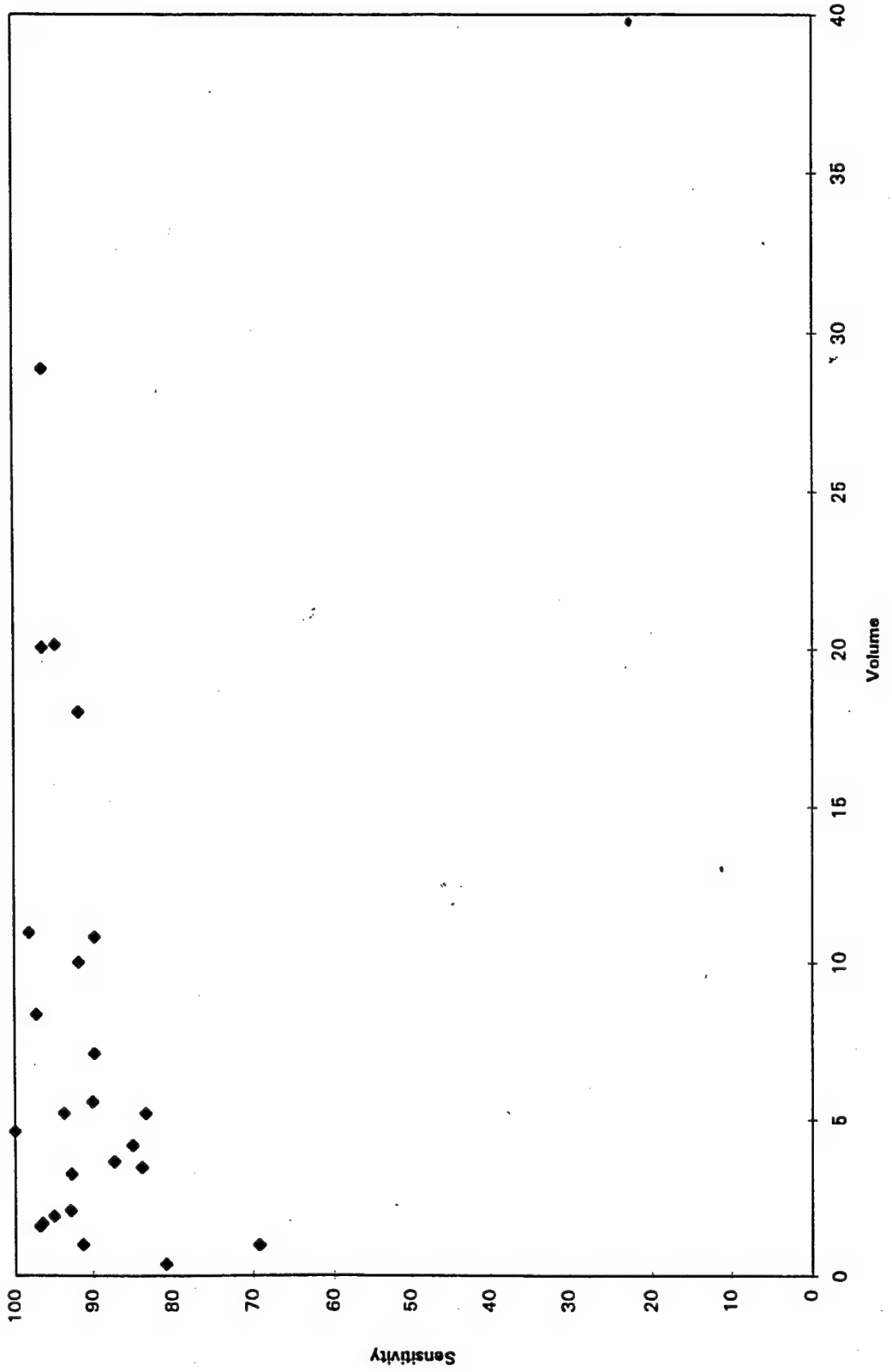
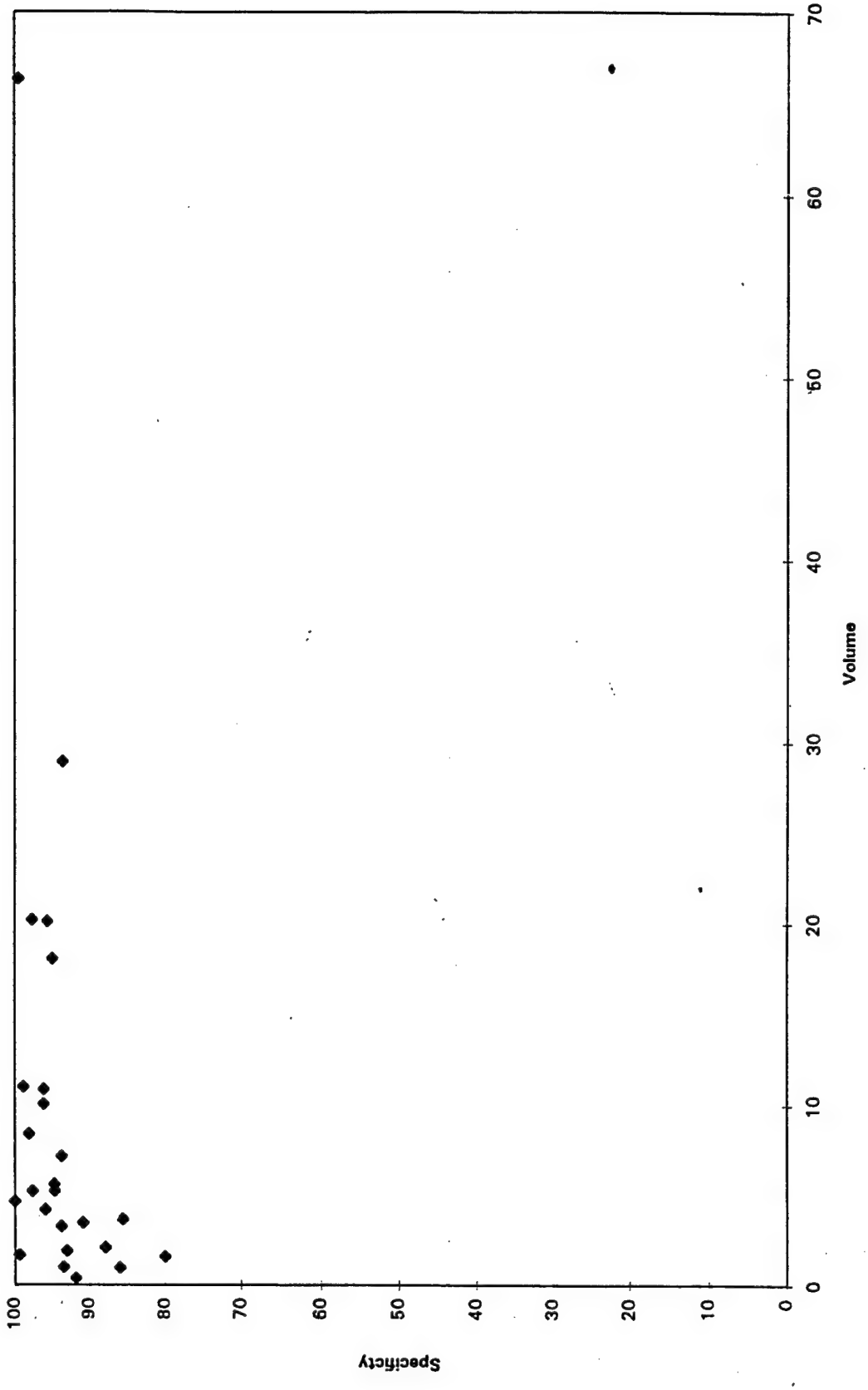


Chart 13

Specificity v. Volume per Operator per Unit of Time



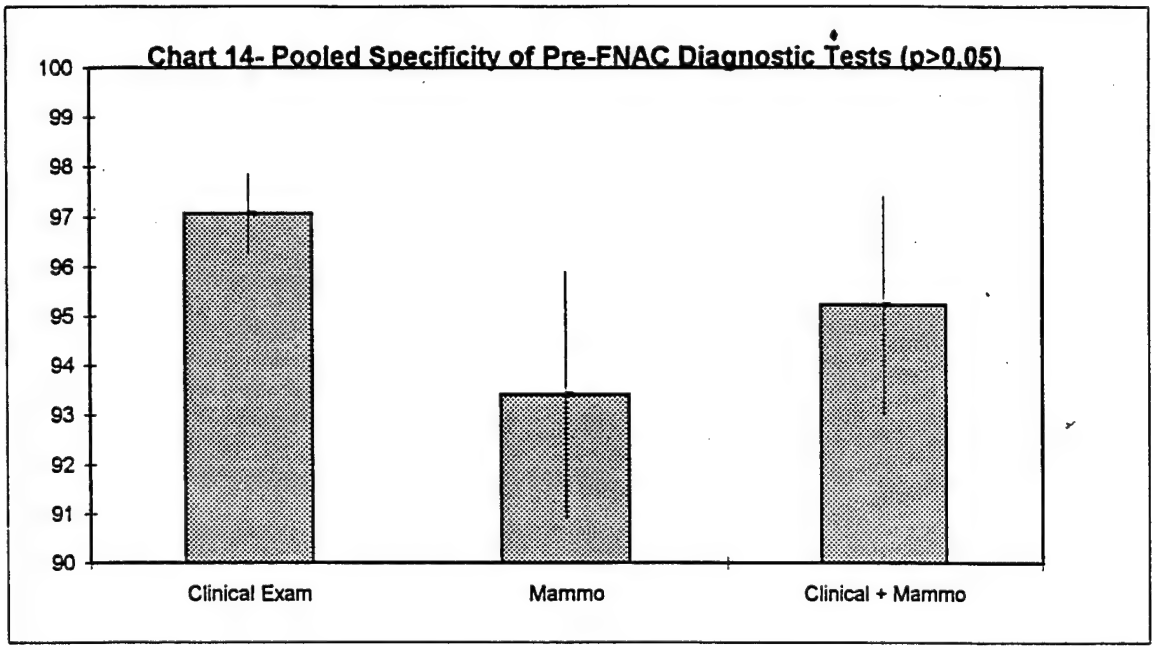


Chart 14- Pooled Specificity of Pre-FNAC Diagnostic Tests (p>0.05)
100
99
98
97
96
95
94
93
92
91
90
Clinical Exam
Mammo
Clinical + Mammo

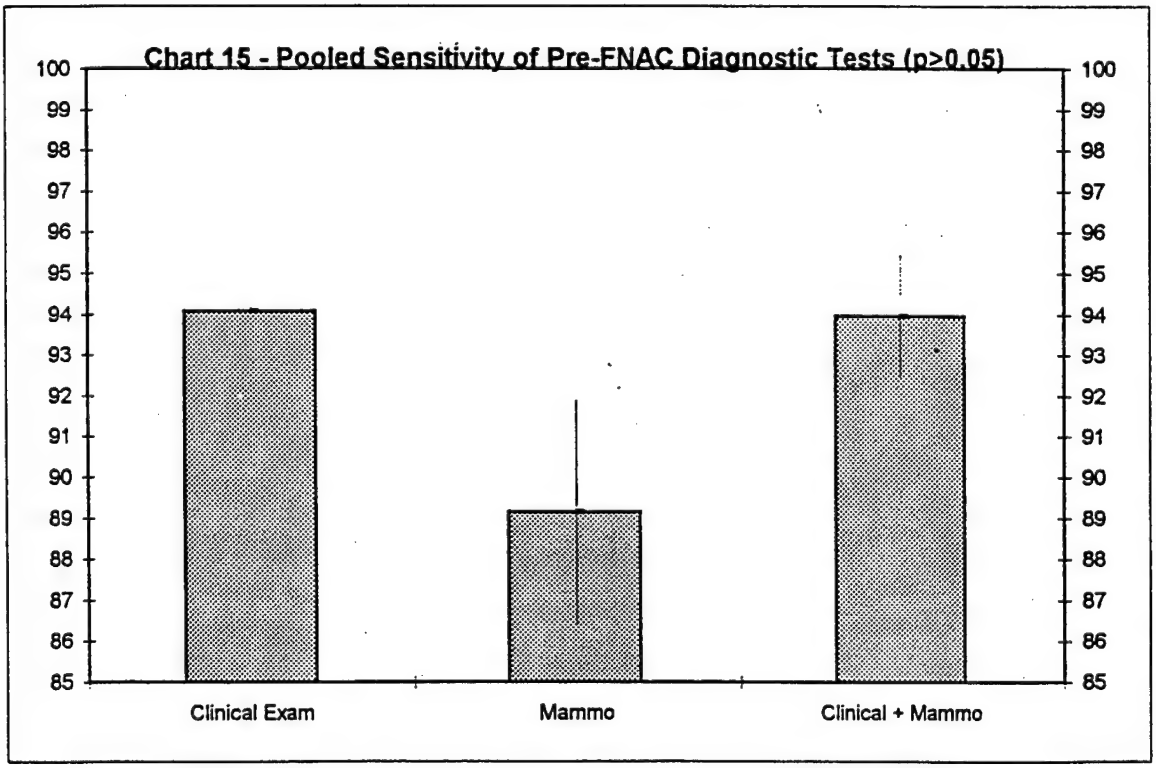


Chart 15 - Pooled Sensitivity of Pre-FNAC Diagnostic Tests (p>0.05)
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
Clinical Exam
Mammo
Clinical + Mammo

Chart 16 - Pooled Specificity for Palpable v. Stereotactic Localization Modes (p>0.05)

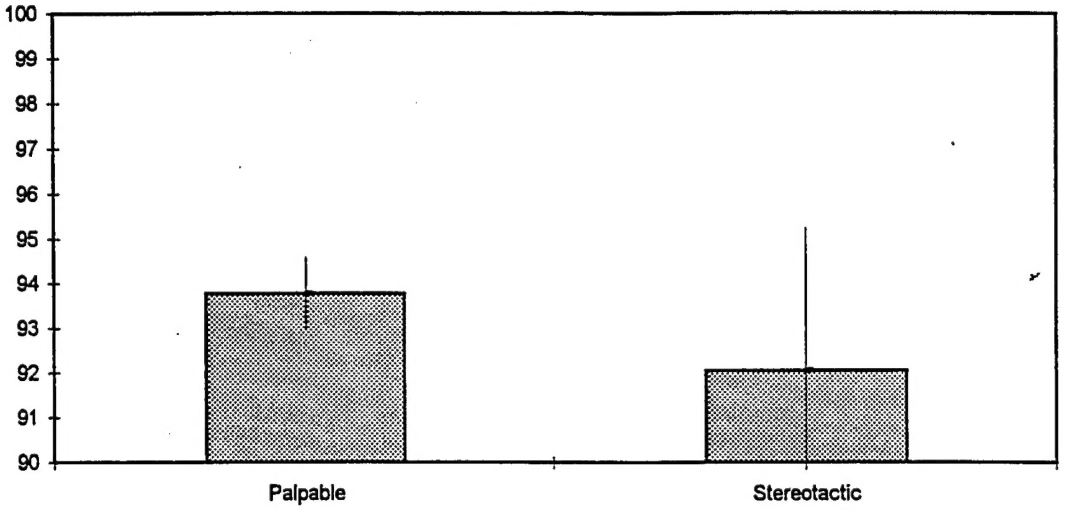


Chart 17 - Pooled Sensitivity for Palpable v. Stereotactic Localization Modes (p>0.05)

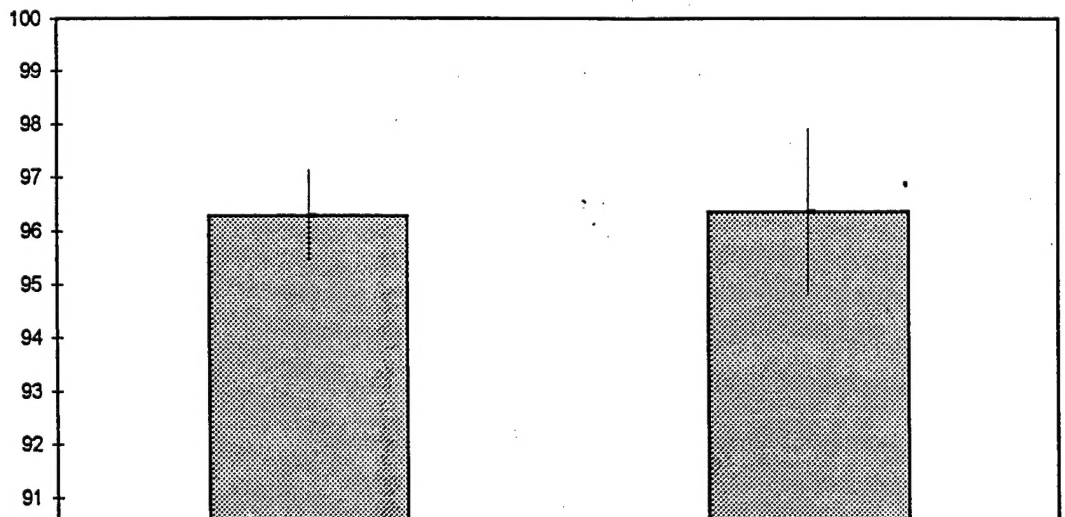


Chart 18

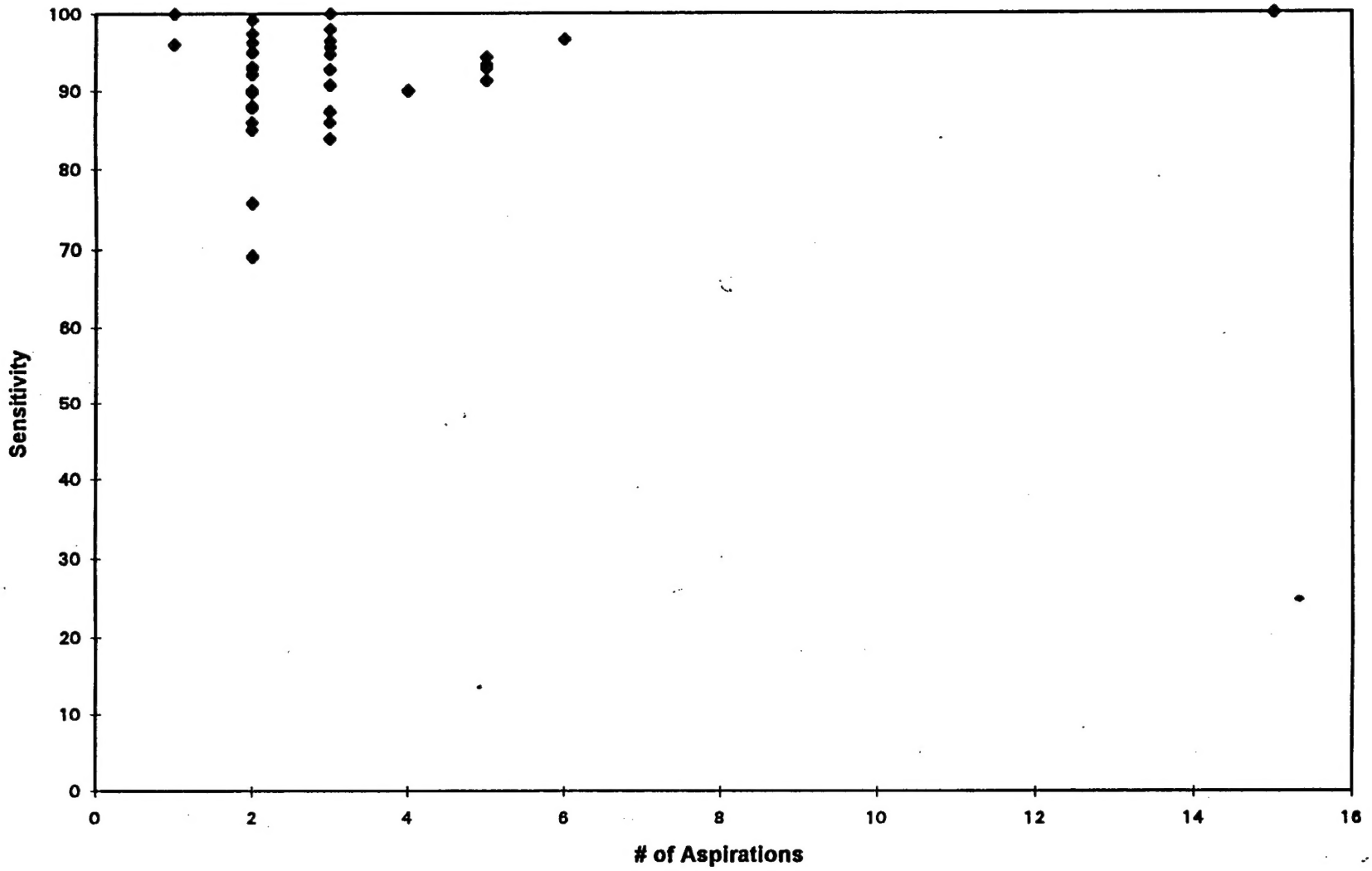

Sensitivity v. # of Aspirations Performed
Sensitivity
100
90
80
70
60
50
40
30
20
10
0
0
2
4
6
8
10
12
14
16
# of Aspirations

Chart 19

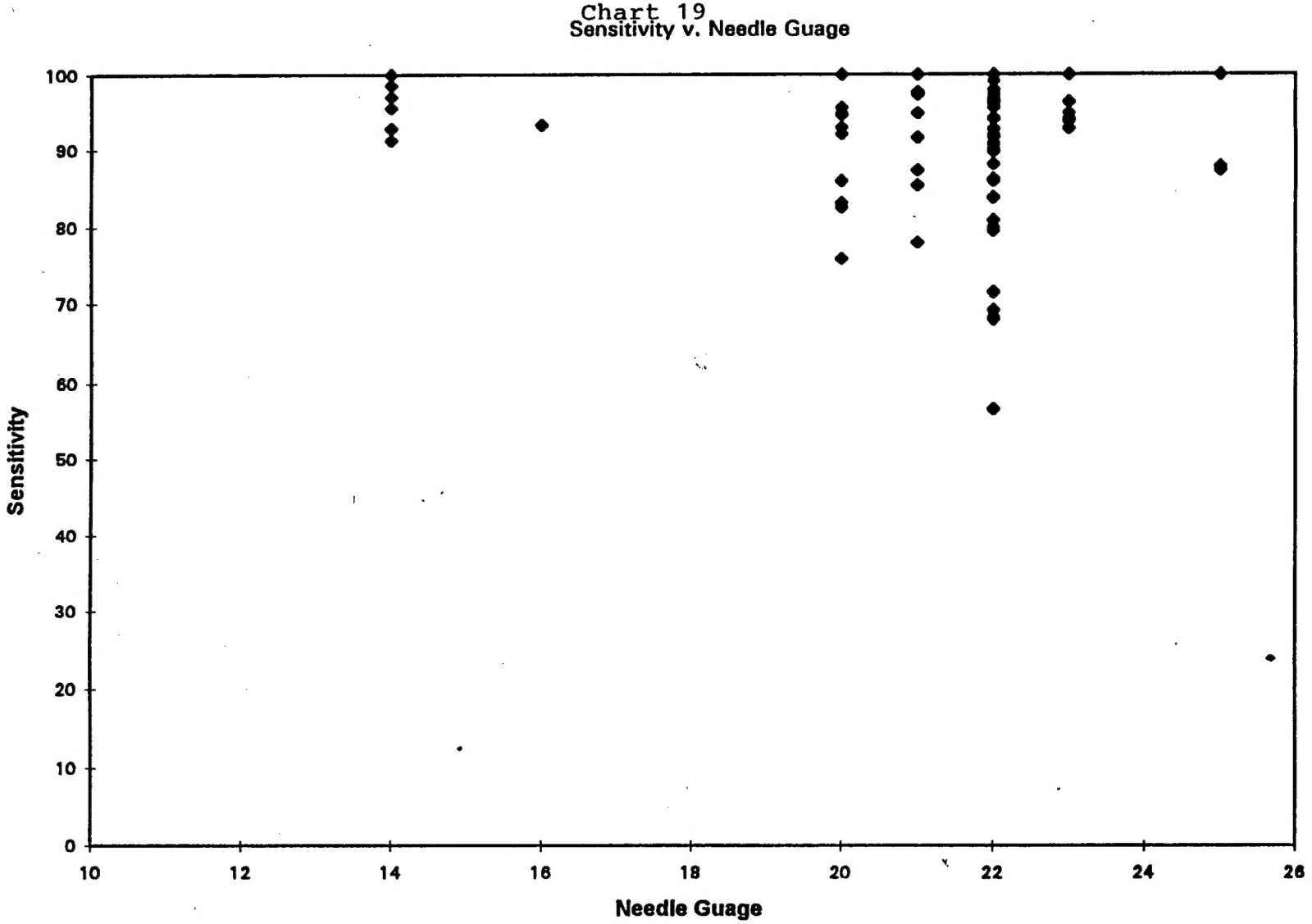

Sensitivity v. Needle Guage
Sensitivity
100
90
80
70
60
50
40
30
20
10
0
10
12
14
16
18
20
22
24
26
Needle Guage

Chart 20

Sensitivity v. Date of Publication